

**MOLECULAR STUDIES ON HEREDITARY
HAEMORRHAGIC TELANGIECTASIA FAMILIES
WITH PULMONARY ARTERIOVENOUS
MALFORMATIONS**

By

Gillian Margaret Fleming Wallace

MBChB , MRCP

Thesis submitted for Degree of
Doctor of Medicine

The University of Edinburgh

2000



CONTENTS

	Page
CONTENTS	I
PUBLICATIONS ARISING FROM THIS THESIS	IX
ABSTRACT	X
LIST OF TABLES	XII
LIST OF FIGURES	XIII
ABBREVIATIONS	XV
DECLARATION	XVII
ACKNOWLEDGEMENTS	XVIII
1. INTRODUCTION	1
1.1. Clinical	2
1.1.1. History	2
1.1.2. Epidemiology	3
1.1.3. Vascular structure	4
1.1.4. Clinical features	4
1.1.4.1. Nose	4
1.1.4.2. Skin and buccal mucosa	5
1.1.4.3. Gastrointestinal tract	6
1.1.4.4. Hepatic	7
1.1.4.5. Urinary system	8
1.1.4.6. Cerebral	8

1.2. Pulmonary arteriovenous malformations	11
1.2.1. Epidemiology and presentation	11
1.2.2. Diagnosis	12
1.2.2.1. Oxygen saturation	13
1.2.2.2. Chest X-ray	14
1.2.2.3. Thoracic CT scanning	15
1.2.2.4. Intrapulmonary shunt investigations	15
1.2.2.5. Pulmonary angiography	16
1.2.3. Treatment	17
1.2.3.1. Embolisation	17
1.2.3.2. Surgery	18
1.2.3.3. Antibiotics	19
1.2.4. Pregnancy and oestrogen	19
1.2.4.1. Implications for management of HHT patients	20
1.3. Genetics of HHT	22
1.3.1. Phenotype-genotype correlations	22
1.3.2. <i>Endoglin</i>	23
1.3.3. <i>Endoglin</i> promoter	25
1.3.4. <i>ALK1</i>	28
1.4. TGF β : the role of endoglin and ALK1 in the superfamily	30
1.5. Linkage analysis	34
1.5.1. Alleles and coinheritance	34
1.5.2. Polymorphic markers	34
1.5.3. Genetic distances	35

1.5.4. Lod scores	36
1.5.5. Potential pitfalls in linkage analysis	37
1.5.5.1. Misphenotyping	37
1.5.5.2. Non parentage	38
1.5.5.3. Inherent limitations in linkage studies	39
1.5.5.3.1. Marker distribution	39
1.5.5.3.2. Reduced allele informativeness	40
1.5.5.3.3. Genome instability	41
1.5.6. Linkage studies versus association studies	41
1.6. Introduction to experimental work	44
2. MATERIALS AND METHODS	52
2.1. Clinical	53
2.1.1. Acquisition of subjects	53
2.1.1.1. Linkage analyses	53
2.1.1.2. Promoter analyses	53
2.1.2. Clinical evaluation	54
2.1.2.1. Screening for PAVMs	54
2.1.2.2. Accurate phenotyping	55
2.1.3. Designation of disease status	56
2.1.4. Sample collection	56
2.2. DNA extraction	57
2.2.1. Genomic DNA	57
2.2.2. Plasmid DNA	59
2.2.3. DNA extraction from an agarose gel	60

2.3. Preparation of leucocytes and plasma	62
2.4. DNA analysis	63
2.4.1. Gel electrophoresis	63
2.4.1.1. Agarose	63
2.4.1.2. Denaturing polyacrylamide	63
2.4.2. Detection systems	64
2.5. Enzyme manipulations of nucleic acids	65
2.5.1. Restriction endonuclease digests	65
2.5.2. Ligations	65
2.5.3. Labelling of PCR products	66
2.6. DNA amplification: polymerase chain reaction	67
2.6.1. Taq polymerase-based PCR	67
2.6.2. Tth/vent polymerase-based PCR	68
2.6.3. Primers	69
2.7. DNA amplification: cloning	71
2.7.1. Insert preparation and ligation	71
2.7.2. Transformation	71
2.8. Sequencing	73
2.9. Linkage	74
2.9.1. Ascertainment of phenotype and genotype	74
2.9.2. Linkage analyses	74
2.9.2.1. Multipoint analyses	75
2.9.2.2. Haplotype analyses	75
3. ENDOGLIN PROMOTER	76

3.1. Amplification of sequence	77
3.2. Sequencing	79
3.3. Fortuitous outcome of primer GW8 misannealing	81
3.4. Familial segregation of polymorphism	83
3.4.1. Restriction endonuclease digests to analyse family D	83
3.4.1.1. XbaI results	83
3.4.1.2. BsrI results	84
3.4.2. Cycle sequencing to analyse family D	84
3.5. Transcription factor binding sites disrupted by the CA deletion	85
3.6. Allele frequency of the polymorphism	86
3.7. Summary	87
4. LINKAGE	101
4.1. Family S	102
4.1.1. Results of clinical evaluation	102
4.1.1.1. Family S members with a diagnosis of PAVMs	102
4.1.1.2. Further details of family S members with a diagnosis of PAVMs	103
4.1.2. Linkage analyses	105
4.1.2.1. Exclusion of <i>endoglin</i>	105
4.1.2.2. Exclusion of <i>ALK1</i>	105
4.1.3. Allele frequencies	106
4.1.4. Haplotype analyses	106
4.1.5. Summary	107
4.2. Family P	108

4.2.1. Results of clinical evaluation	108
4.2.1.1. Family P members with possible PAVMs	109
4.2.2. Linkage analyses	110
4.2.2.1. Inconclusive <i>endoglin</i> linkage data	110
4.2.2.1.1. Haplotype analyses	111
4.2.2.1.2. Models constructed for inconclusive <i>endoglin</i> linkage data	111
4.2.2.1.2.1. Model 1: Analyses have failed to confirm linkage in an <i>endoglin</i> -linked family	111
4.2.2.1.2.2. Model 2: In family P there is a novel HHT disease gene sited approximately 20-30cM away from the <i>endoglin</i> locus	113
4.2.2.1.2.3. Model 3: There is no HHT disease gene in the region on chromosome 9. The <i>D9S61</i> and <i>HXB</i> lod scores are spuriously high as a result of common alleles present in the population apparently segregating with disease	115
4.2.2.1.2.4. Conclusions from the models	116
4.2.2.2. Exclusion of <i>ALK1</i>	116
4.2.3. Allele frequencies	117
4.2.4. Summary	117
4.3. Families V and T- assessment of <i>ALK1</i> linkage in known non- <i>endoglin</i> HHT-PAVM families	118
4.3.1. Family V	119

4.3.2. Family T	119
4.3.3. Summary	119
4.4. Conclusions	120
5. DISCUSSION	141
5.1. Clinical issues	142
5.1.1. Pregnancy	142
5.1.1.1. PAVM rescreening	142
5.1.2. Microscopic PAVMs	143
5.2. Molecular issues	145
5.2.1. The importance of finding a polymorphism close to the <i>endoglin</i> promoter	145
5.2.1.1. Should the <i>endoglin</i> promoter be screened for HHT mutations?	145
5.2.1.2. Functional effects of any promoter polymorphism found	146
5.2.1.3. Other regions to be screened	147
5.2.2. The importance of identifying the third gene	147
5.2.2.1. Elucidation of the biochemical role	149
5.2.3. Phenotype/ Genotype correlations	150
5.3. Vascular biology issues	152
5.3.1. Why are some vessels abnormal and not others?	152
5.4. Concluding remarks	153
5.5. Further work	154
5.5.1. Clinical	154
5.5.2. Molecular	154

5.5.3. Vascular biology	155
BIBLIOGRAPHY	156
APPENDIX A – CLINICAL QUESTIONNAIRE	177
APPENDIX B - RECIPES	178
APPENDIX C – PRIMERS FOR THE <i>ENDOGLIN</i> PROMOTER	180
APPENDIX D – PRIMERS FOR <i>ENDOGLIN</i> LINKAGE	181
APPENDIX E – PRIMERS FOR <i>ALK1</i> LINKAGE	182

PUBLICATIONS ARISING FROM THIS THESIS

Chapter 1

Wallace GMF, Jackson JE, Hughes JMB, Shovlin CL

The risk of cerebral haemorrhage in families with HHT and pulmonary AVMs.

Thorax 1999; 54(suppl 3): A32.(Abstract)

Easey A, **Wallace GMF**, Jackson JE, Taylor W, Shovlin CL

Retrospective analysis of the risk of cerebral bleeding in hereditary haemorrhagic telangiectasia.

IVth HHT Scientific Conference-Conference programme and abstracts April 2001:

79

(Abstract) Full manuscript to be submitted in December 2001

Chapter 3

Wallace GMF, Bernabeu C, Shovlin CL

Analysis of the promoter region of the endoglin gene in families with hereditary haemorrhagic telangiectasia and pulmonary arteriovenous malformations.

American Journal of Respiratory and Critical Care Medicine 1999; 159(3): A351.

(Abstract)

Chapter 4

Wallace GMF, Shovlin CL

An HHT family with pulmonary involvement is unlinked to the known HHT genes, endoglin and ALK1.

American Journal of Respiratory and Critical Care Medicine 2000; 161(3): A641.

(Abstract).

Wallace GMF, Shovlin CL

A hereditary haemorrhagic telangiectasia family with pulmonary involvement is unlinked to the known HHT genes, endoglin and ALK1.

Thorax 2000; 55(8): 685-90.

Cole SG, Begbie ME, Chaudhary M, **Wallace GMF**, Shovlin CL

Linkage studies on a PAVM-HHT family unlinked to endoglin or ALK-1.

IVth HHT Scientific Conference-Conference programme and abstracts April 2001:

74

(Abstract)

ABSTRACT

This thesis comprises two separate molecular studies on Hereditary Haemorrhagic Telangiectasia (HHT) families with pulmonary arteriovenous malformations (PAVMs).

PAVMs occur in over 25% of patients with the autosomal dominant disorder HHT. They account for some of the most devastating consequences of the disease.

Mutations in two genes, *endoglin* and *ALK1*, are known to cause HHT. Both encode a protein expressed on vascular epithelial cells and are involved in signalling by members of the transforming growth factor (TGF) β superfamily. To date, PAVMs have not been detected in *ALK1* families. It has been suggested therefore that clinical screening for PAVMs is restricted to *endoglin*-linked families. There is evidence from a single HHT family without pulmonary involvement that a third HHT gene may exist.

In one study in this thesis the *endoglin* promoter region was screened for mutations in four endoglin-linked families, and four other small families, in whom no mutations had previously been found studying all coding regions and splice sites of the gene. No mutations were found. However a two base pair CA deletion polymorphism, with an allelic frequency of 9% in the UK population, was found 2563 base pairs upstream of one of the major *endoglin* transcription start sites. It may prove a useful tool for association studies.

In the other study linkage analyses were performed on four PAVM-HHT families, two of which were expanded during this study. They had all previously been found

to be unlinked to *endoglin*. Linkage to both *endoglin* and *ALK1* was significantly excluded in one family. Significant linkage to *ALK1* was found in another. It is concluded that a third HHT gene exists and that PAVMs can occur in all genotypes with resulting clinical implications.

Confirmation by this work that at least a third HHT gene exists should precipitate identification of the gene and elucidation of its biochemical role. It seems reasonable to speculate that it will encode another component of the TGF β signalling complex or a downstream effector, and so its identification should increase our understanding of this superfamily of signalling molecules. Further work however is still required to determine how the actions of the TGF β signalling complex lead to abnormal vascular structures.

LIST OF TABLES

	Page
Table 4.1 Clinical features of family S	121
Table 4.2 Diagnoses of HHT	122
Table 4.3 Analysis of lod scores with equal and adjusted allele frequencies in family S	123
Table 4.4 Clinical features of family P	124
Table 4.5 Two point lod scores for family P with markers <i>D9S60</i> , <i>D9S315</i> , and <i>D9S61</i> flanking the <i>endoglin</i> locus	126
Table 4.6 Two point lod scores for family P with additional chromosome 9 markers	127
Table 4.7 Two point lod scores for markers adjacent to <i>ALK1</i> in family P	128
Table 4.8 Calculating allele frequencies of marker <i>D9S315</i> in family P	129
Table 4.9 Analysis of lod scores with equal and adjusted allele frequencies in family P	130
Table 4.10 Two point lod scores for markers adjacent to <i>ALK1</i> in family V and family T	131

LIST OF FIGURES

	Page
Figure 1.1 Histology of microscopic PAVMs	46
Figure 1.2 Telangiectasia of the lips	47
Figure 1.3 CT head scan showing CAVM	48
Figure 1.4 Chest X-ray showing PAVM	49
Figure 1.5 Chest CT scan showing PAVM	50
Figure 1.6 Pulmonary angiogram pre and post embolisation	51
Figure 3.1 Region around the <i>endoglin</i> promoter and the position of primers used	88
Figure 3.2 PCR products using primers GW1 and GW7 and genomic target DNA	89
Figure 3.3 Restriction endonuclease digest with <i>Ava</i> I and <i>Pvu</i> II of PCR product using primers GW1 and GW7 on cloned genomic <i>endoglin</i>	90
Figure 3.4 PCR products using primers GW1 and GW3 and genomic <i>endoglin</i>	91
Figure 3.5 Comparison of GenBank sequences AF035753 and X72012	92
Figure 3.6 The G track of sequence using primer GW8	93
Figure 3.7 Colonies digested with <i>Eco</i> RI to obtain cloned GW1/10 insert	94
Figure 3.8 The G track of 10 sequenced clones	95
Figure 3.9 Sequencing gels showing G and T bases absent in clone 2	96
Figure 3.10 Family tree of family D	97

Figure 3.11 Restriction endonuclease digest with XbaI and BsrI of PCR product using primers GW5 and 10 on DNA from members of family D	98
Figure 3.12 Family D sequencing gel	99
Figure 3.13 Terminations in G track sequencing with GW16 on GW5/10 PCR product to assess allele frequency of polymorphism	100
Figure 4.1 Family tree of family S	132
Figure 4.2 The position of <i>endoglin</i> and its flanking markers	133
Figure 4.3 Two point and multipoint lod scores spanning the <i>endoglin</i> locus in family S	134
Figure 4.4 The position of <i>ALK1</i> and its flanking markers	135
Figure 4.5 Two point and multipoint lod scores spanning the <i>ALK1</i> locus in family S	136
Figure 4.6 Limited haplotype analyses of chromosome 9 and 12 markers in family S	137
Figure 4.7 Family tree of family P	138
Figure 4.8 Illustration of allele inheritance of chromosome 9 markers in family P	139
Figure 4.9 Family tree of family V and family T	140

ABBREVIATIONS

A	Adenine
ABG	Arterial Blood Gas
ALK1	Activin Receptor-Like Kinase 1
AVM	Arteriovenous Malformation
BAEC	Bovine Aortic Endothelial Cells
BSA	Bovine Serum Albumin
C	Cytosine
CAVM	Cerebral Arteriovenous Malformation
CLS	Dr Claire Shovlin
CT scan	Computed Tomographic scan
CVA	Cerebrovascular Accident
CXR	Chest X-ray
DMSO	Dimethylsulphoxide
DNA	Deoxyribose Nucleic Acid
EDTA	Ethylenediaminetetraacetate
FBC	Full Blood Count
FEV1	Forced Expiratory Volume (in 1 second)
FVC	Forced Vital Capacity
G	Guanine
GI	Gastrointestinal
GKLF	Gut-Enriched Kruppel-Like Factor
GTE	Genomic Tris EDTA

HGMP	Human Genome Mapping Project
HHT	Hereditary Haemorrhagic Telangiectasia
IPTG	Isopropylthio- β -D-Galactoside
KCO	Transfer coefficient
MREC	Multicentre Research Ethics Committee of Scotland
MRI	Magnetic Resonance Imaging
NCBI	National Center for Biotechnology Information
PA	Posteroanterior
PAI	Plasminogen Activator Inhibitor
PAVM	Pulmonary Arteriovenous Malformation
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
RNA	Ribonucleic Acid
SDS	Sodium Dodecyl Sulfate
SNP	Single Nucleotide Polymorphism
T	Thymidine
TAE	Tris Acetate EDTA
TBE	Tris Borate EDTA
TE	Tris-EDTA
TGF β	Transforming Growth Factor β
TLCO	Transfer Factor

DECLARATION

I declare that the work contained in this thesis is original. I have been solely responsible for the organisation and day-to-day running of the studies contained herein, as well as all aspects of data collection and the analysis of results, unless otherwise referenced.

Gillian M.F. Wallace

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank those people who have been invaluable to this thesis in terms of their support, encouragement, and guidance over the past two years.

Firstly I would like to thank Dr Claire Shovlin (Respiratory Unit, Hammersmith Hospital, London and previously Respiratory Unit/ Rayne Laboratories, Edinburgh) for her expert knowledge and advice with this work.

I would also like to acknowledge Professor Chris Haslett and Professor Jonathan Lamb (Rayne Laboratories, Edinburgh) for letting me work in their laboratories.

I would like to thank Dr Karen Tan, Dr Grainne Cunningham, Mr Mark Marsden, Dr Marie-Helene Ruchard-Sparagano and other members of the Rayne laboratories for their helpful advice.

I should also like to thank the British Heart Foundation for funding this work.

I would like to thank my family and friends for their support. I am particularly indebted to my husband, Dr Andrew Riley, for his enormous help with this work throughout the two years, both with encouragement and with greatly appreciated computing skills.

1. INTRODUCTION

The classical patient with HHT has epistaxis, telangiectasia around the lips and finger tips, and a similarly affected relative. However several other organs may also be involved accounting for some of the more devastating clinical complications of the disease. The manifestations are all due to abnormalities of vascular structure.

Mutations in two genes, *endoglin* and *ALK1*, are known to cause HHT. They both encode cell-surface proteins involved in signalling by the TGF β superfamily. To date, PAVMs have not been detected in *ALK1*-linked families. There is evidence from one family with no pulmonary involvement that a third gene may exist.

1.1. CLINICAL

1.1.1. HISTORY

Recurrent epistaxis, often associated with haemoptysis and occurring in several family members, was first noted by Sutton in 1864¹. A hereditary tendency to epistaxis was confirmed one year later by Babington². The first report of skin lesions being associated with hereditary epistaxis was made in 1876 by Legg. His contribution was overlooked, as he suggested that the features were associated with haemophilia and described the skin lesions too imprecisely as “naevi”³. The naevi were redefined as telangiectases by Chiari in 1887 but haemophilia continued to be diagnosed⁴. Rendu, in 1896, distinguished this disease from haemophilia, by the lack of haemorrhage from skin lacerations and tooth extractions⁵. He also discussed associated gastrointestinal bleeding. The disease was brought to proper attention by Osler in 1901 who gave a classic description of the characteristic triad of recurrent epistaxis, hereditary pattern, and multiple cutaneous telangiectases⁶. This provoked several other reports of telangiectasia with epistaxis⁷⁻¹⁰ but it was Weber in 1907 who most clearly further defined its clinical manifestations and behaviour¹¹ leading, possibly unfairly, to the eponymous name Rendu-Osler-Weber syndrome. Hanes, in 1909, most clearly described the inheritance pattern of the disease, described the appearance of telangiectases under the light microscope, and gave the disease the name Hereditary Haemorrhagic Telangiectasia (HHT)¹².

The existence of pulmonary capillaries was first described in 1661 by Malpighus¹³. Churton in 1897 described abnormalities of the pulmonary vasculature in the form of

“aneurysms as large as walnuts” in a boy who had haemoptysis and epistaxis¹⁴.

Pulmonary arteriovenous fistula was described in 1939 by Smith and Horton¹⁵.

Rundles in 1945 described a pulmonary artery aneurysm in an HHT patient who also had gastric telangiectasia¹⁶. Gastric telangiectasia had been previously described by Renshaw in 1936¹⁷. In 1947, Whitaker postulated that cyanosis associated with “cavernous haemangioma of the lung” is due to its shunt effect¹⁸.

1.1.2. EPIDEMIOLOGY

HHT occurs with a wide geographic distribution among many ethnic and racial groups including negro Americans^{19 20}, Mexicans²¹, Chinese²², Indians²³, Thai²⁴, and Japanese²⁵. Studies of prevalence give varying figures of 1 in 2351 in the French region of Ain²⁶, 1 in 3500 on the Danish island of Funen²⁷, 1 in 5155 in the Dutch Leeward Islands²⁸, 1 in 16,500 in Vermont²⁹, and 1 in 39,216 in northern England although this is thought to be almost certainly an underestimate³⁰. These figures, and others recently published³¹, show that, at least in the population investigated, HHT is more prevalent than previously thought, at around 1 in 10,000. Figures of around 1 in 100,000 had been formerly reported^{32 33}.

The disease is autosomal dominant with an estimated penetrance overall of 97%^{30 34}. This, however, is age-related with manifestations developing throughout life^{34 35} and varying between affected individuals even from the same family^{36 37}. There are very few reports of homozygous cases^{38 39} in one of whom the condition was fatal. Thus affected individuals are almost exclusively heterozygous.

1.1.3. VASCULAR STRUCTURE

An AVM consists of a dilated aneurysmal sac of arterial and venous origin supplied by one or more feeding arteries and draining to one or more dilated veins. There is no intervening capillary bed⁴⁰(see Figure 1.1). 95% of PAVMs derive their arterial supply from the pulmonary rather than the systemic circulation⁴¹. Telangiectases are also abnormal vascular structures but instead are due to dilated arterioles, capillaries and postcapillary venules⁴²⁻⁴⁴.

1.1.4. CLINICAL FEATURES

1.1.4.1. NOSE

Spontaneous, recurrent nose bleeds from telangiectases of the nasal mucosa is the most common clinical manifestation of HHT. It occurs in more than 90% of affected individuals⁴⁵ beginning by the age of 10 years in many and the age of 21 years in most⁴⁵. Controversy exists regarding the progression of epistaxis with age. Some⁴⁶ have demonstrated that it remains stable over time, McCaffrey et al have shown a group with 25% progressive, 45% static, and 17% regressive symptoms⁴⁷, and some have found that it tends to be progressive in the majority^{34 45 48}. The severity is very variable between individuals. A mortality rate of up to 4% has been quoted⁴⁹ although no recent data is available.

Many require no treatment. Others may require nasal packing, multiple transfusions, and oral iron supplementation. Many other more definitive treatments have been tried. Argon and ND Yag laser treatment has a beneficial effect but data on long term effects is scarce^{50 51}. Septal dermoplasty in adept hands has good results in

patients with severe epistaxis but vessels usually regrow^{52 53}. Oral high dose oestrogen therapy (1-5mg oestradiol/day) has been tried. Oestrogens appear to be effective in uncontrolled studies⁴⁸ but controlled studies do not support efficacy^{54 55}. The possible complications of long term oestrogen must also be considered⁴⁸. Nevertheless oestrogen and progesterone receptors have been shown to be present in nasal telangiectases in females (progesterone receptors only in males) compared to none in normal controls⁵⁶. The mechanism of any possible action remains unclear, but it has been suggested that oestrogens improve continuity of telangiectatic endothelium and so decrease bleeding tendency⁴³, or they may act by inducing squamous metaplasia⁴⁸. Transcatheter embolotherapy of arteries leading to the nasal mucosa has been tried. Older studies indicate it may be less effective in HHT patients than other individuals⁵⁷⁻⁵⁹, but there are no data regarding results using more modern methods.

1.1.4.2. SKIN AND BUCCAL MUCOSA

Telangiectases of the skin and buccal mucosa occur in about 75% of individuals³⁴ (see Figure 1.2) typically presenting later in life than epistaxis³⁴ from about the third decade of life⁴⁹ and increasing in size and number with age^{60 61}. They mostly occur on the face, lips, tongue and buccal mucosa, and fingertips, but can occur elsewhere. They may bleed (bleeding from conjunctival telangiectasia causes “bloody tears” but this is rarely clinically important⁶²). The main concern is cosmetic. Laser ablation can be used⁶³.

1.1.4.3. GASTROINTESTINAL TRACT

Recurrent haemorrhage of the gastrointestinal (GI) tract occurs in a substantial minority (up to 44%) of HHT affected individuals^{27 34 64} and is the second most frequent site of haemorrhage⁶⁴. It often presents as an iron deficiency anaemia but occasionally as an acute GI haemorrhage²⁷. The onset is usually from the fifth or sixth decade⁶⁴. In half the patients the site of bleeding is telangiectases in the stomach or duodenum. Far fewer have blood loss due to colonic telangiectasia^{27 64 65}. They are visualised by endoscopy and are similar in size and appearance to mucocutaneous telangiectases but may be surrounded by an anaemic halo²⁷. Less commonly arteriovenous malformations (AVMs) and aneurysms may occur, depicted by GI angiography^{64 66}. Most patients are satisfactorily managed conservatively with, if necessary, blood transfusions or oral iron therapy⁶⁴. However laser therapy may also be used to control bleeding in the short term, but less commonly in the long term^{65 67}. Recurrence of bleeding is very common, due to missed or incompletely treated lesions or recurrent disease⁶⁷. Immediate or delayed haemorrhage from the laser treatment, or perforation, may occur⁶⁸. A low dose combination of oestrogen (0.05mg oestradiol/day (a “high contraceptive dose”)) and progesterone has been shown to be beneficial in controlling severe blood loss through GI telangiectasia, and for the effect to remain for several months after cessation of treatment⁶⁹. The mode of this effect has yet to be determined⁶⁹. Effectiveness of treatment with aminocaproic acid (antifibrinolytic) has been found to be variable^{70 71}. Surgery has limited success due to recurrent disease⁶⁷.

1.1.4.4. HEPATIC

Hepatic involvement may occur in up to 30% of HHT patients⁷². Large arteriovenous malformations (AVMs) may occur between the hepatic artery and vein and can result in a substantial left to right shunt^{73 74} proven by a higher oxygen saturation in the hepatic vein than the inferior vena cava⁷⁴. This may give rise to high output cardiac failure⁷⁴. Shunts between the hepatic artery and portal vein occur more rarely and lead to hepatic encephalopathy⁷⁵ particularly after GI tract bleeding. They can also lead to portal hypertension with associated oesophageal varices⁷⁶. Ascites has been reported in association with multiple AVMs but no portal hypertension⁷⁷. Spontaneous bleeding from liver AVMs has not been reported. AVMs can occur alone or in combination with fibrosis of the liver⁷⁸. Linkage of fibrous bands results in nodularity of the hepatic parenchyma but without disruption of the architecture within the nodules as in true cirrhosis^{78 79}. This is known as “cirrhosis of HHT”. The clinical significance of this is debated⁷⁹. True cirrhosis can also occur either due to coexisting disease, for instance cardiac failure or alcoholism, or due to treatment for HHT symptoms, for example massive blood transfusion or long term oestrogen treatment.⁸⁰ Hepatic AVMs, suspected by hepatomegaly, a liver bruit, and sometimes raised γ GT and alkaline phosphatase levels⁷⁴ can be diagnosed by angiography, CT, MRI, or doppler sonography^{72 74}⁸¹. Doppler sonography has been suggested to be the ideal method for screening⁷² although screening is not performed in the U.K. Treatment of AVMs is largely conservative⁸². Embolisation of feeding vessels is performed⁷³ but a 25% mortality

rate was calculated in 1996 from all the data available⁸². Liver transplantation may be necessary if disease is extensive⁸³. Following several cases of fatal hepatic necrosis occurring after hepatic embolisation, the United States HHT physicians have placed a moratorium on embolisation of hepatic AVMs where portal shunting cannot be definitely excluded, unless supported by hepatic transplantation⁸⁴.

1.1.4.5. URINARY SYSTEM

Involvement of this system is rare (0.6%)³⁴. Telangiectasia can occur over the bladder and urethra⁸⁵. Treatment other than iron therapy is rarely required⁸⁵.

1.1.4.6. CEREBRAL

5-10% of HHT patients are thought to have cerebral involvement^{34 86} with telangiectases, AVMs (cerebral AVMs (CAVMs)), aneurysms, or cavernous angiomas⁸⁷ (see Figure 1.3). However screening of asymptomatic patients indicates a higher incidence⁸⁶. In a literature review of 77 cases of cerebral involvement, 47% had telangiectases, 22% CAVMs, 8% aneurysms, and 22% spinal vascular malformations⁸⁷. Most complications arise from CAVMs. They can lead to headache, seizures, ischaemia of the surrounding tissue due to a steal effect, or haemorrhage⁸⁷. It is unclear whether the natural course of CAVMs in HHT is the same as for CAVMs unassociated with HHT, in whom the average yearly bleeding risk is between 2 and 3%⁸⁸, but it is usually assumed to be the same. The haemorrhage risk of a CAVM has been found to vary with certain features of the lesion: small size of CAVM⁸⁸, deep venous drainage, feeding by perforators, intranidal aneurysm(s), multiple aneurysms, feeding by the vertebrobasilar system,

and location in the basal ganglia have been found to increase the risk of haemorrhage⁸⁹. Overall a haemorrhage risk of 2% per annum^{89 90} is used as the basis of risk-benefit analyses when screening is contemplated. Analyses suggest that in asymptomatic patients the risks of haemorrhage with expectant treatment outweigh the risk of intervention, particularly in young patients⁹¹. However patient age, specific features of the lesion, and available expertise must be considered for each individual⁹¹. Although screening is offered in USA and Canada, it is not yet offered in Europe and the question of whether it should be remains the subject of intense debate.

In order to justify a prospective study to establish the value of CAVM screening and treatment in asymptomatic HHT patients, particularly in younger groups, in the UK, a series of 80 HHT families identified by at least one individual from each presenting to the Hammersmith Hospital was studied by myself and Dr Shovlin. Full family histories revealed that out of 335 individuals, from these 80 families, 25 (7.5%) suffered catastrophic cerebral haemorrhages (almost all fatal or leaving severe residual impairment) with many of the haemorrhages being due to CAVMs and 7 of the patients being aged between 10 and 22⁹². The proposed study therefore seems appropriate.

The best diagnostic approach varies with the different lesions. Telangiectases can only be detected by MRI⁸⁷. MRI is also the best method for detecting cavernous haemangiomas⁹³. MRI and angiography are more sensitive than CT in detecting CAVMs⁸², but significant procedural morbidity and exposure to radiation and IV

contrast with angiography makes this unsuitable as a screening procedure. IV digital subtraction angiography avoids this but the resolution is low and vertebrobasilar system poorly visualised⁸². Therefore MRI seems to be the best technique for screening⁸². It should be noted however that MRI should not be performed in PAVM patients embolised with non-MRI compatible coils. This accounts for the majority of patients who were embolised before the late 1990s.

The method of treatment depends on site and type of lesion, availability, and local expertise. Surgery, embolisation, and stereotactic radiotherapy can be used alone or in combination⁸². A comparison of series of patients, treated by stereotactic radiotherapy or microsurgery, indicates that stereotactic radiotherapy is less efficacious and has a higher associated morbidity and mortality than microsurgery⁹⁴.

1.2. PULMONARY ARTERIOVENOUS MALFORMATIONS

Further information can be found in an extensive recent review of pulmonary arteriovenous malformations in HHT⁹⁵.

1.2.1. EPIDEMIOLOGY AND PRESENTATION

Involvement of the lungs, in the form of pulmonary arteriovenous malformations (PAVMs), occurs in 15-33% of HHT patients^{60 86 96}. Onset is in the relatively young (20% of PAVMs in those under 20 years, 50% in those under 38 years in one series³⁴). At least 70% of all PAVMs occur in HHT patients and so the possibility of HHT should be considered in anyone found to have a PAVM^{97 98}. This is particularly so if there are multiple lesions which are especially common in HHT patients⁹⁷⁻¹⁰⁰. 70% of PAVMs are localised in the lower lung fields^{97 100 101}. PAVMs usually increase in size^{96 98 101}, particularly if they are multiple¹⁰¹, and very rarely regress⁹⁶. Complications from the PAVMs are more common in HHT than non HHT patients^{99 102}. Mortality rates for untreated PAVM patients range from 4-22%^{99-101 103}. This may be due to PAVMs haemorrhaging into a bronchus or the pleural cavity¹⁰⁴⁻¹⁰⁶ causing haemoptysis (11% on average) and haemothorax (less than 1% on average) respectively⁹⁵. The more common complications of PAVMs however arise from the right to left shunts resulting from the capillary-free communications between pulmonary and systemic circulations. Critically, paradoxical embolism can occur through these shunts due to material which is normally trapped by the small diameter capillaries being able to pass through the wider PAVMs. This may lead to transient ischaemic attacks (TIAs) or

cerebrovascular accidents (CVAs) in, on average, 24%, and cerebral abscesses in, on average, 9% of PAVM patients⁹⁵. More rarely other systemic emboli may occur for instance to the kidneys or knee joints¹⁰⁷. PAVMs in fact account for two thirds of all neurological symptoms in HHT patients⁹⁷. A high incidence of migraine, in 43% of HHT patients with PAVMs, has been reported but the underlying reason not elucidated⁹⁷. This has not however been found by others and the data presented in this thesis does not support this finding (see Sections 4.1.1 and 4.2.1). Hypoxaemia occurs due to the shunted blood avoiding oxygenation. It is often well tolerated due to generation of a high cardiac output, facilitated by low pulmonary vascular resistance¹⁰⁸. It does, however, present as dyspnoea on exertion in, on average 47%, and chest pain (due to associated polycythaemia as well as hypoxia) in 12%⁹⁵. Central cyanosis, clubbing, or a vascular bruit are present in, on average, 30%, 36%, and 49% respectively of PAVM patients⁹⁵. Up to 50% of patients with these signs or an abnormal chest x-ray have no respiratory symptoms⁹⁵ and for many of them the first presentation of lung involvement is with a neurological event^{99 100}. 13% of PAVM patients in fact die from fatal CVAs^{99 100}. As treatment for PAVMs is successful¹⁰⁹ screening for them prior to these catastrophic events is vital.

1.2.2. DIAGNOSIS

It is recommended that PAVM screening is performed every 5-10 years^{36 82} or around pregnancy and puberty since these are associated with PAVM enlargement^{106 110}. All genotypes should be screened (see Section 5.2.3). Current recommendations for investigation of PAVMs vary for individuals suspected of

having them through clinical features or screening investigation. One recommendation is to perform a PA and lateral CXR, arterial blood gas (ABG) analysis and/or supine and erect oximetry. If these raise concern at least one non-invasive further investigation is recommended before pulmonary angiography is considered⁹⁵. Alternative strategies used in different institutions include use of contrast echocardiography as an initial screen¹¹¹, arterial blood oxygen levels rather than oximetry⁸⁶, measurement of the right to left shunt breathing 100% oxygen⁸⁶, and more extensive use of CT scanning¹¹².

1.2.2.1. OXYGEN SATURATION

As mentioned in Section 1.2.1, PAVMs lead to hypoxaemia due to the anatomic right-to-left shunt provided by the abnormal vessels which bypass the pulmonary capillary bed. The simplest method for detecting a right-to-left shunt is the demonstration of a low partial pressure of oxygen (PaO₂) in arterial blood by blood gas analysis. The detection of hypoxaemia alone using age-defined PaO₂ [$104\text{mmHg} - 0.24 \times \text{age}$] has been used as a single diagnostic test⁸⁶, but the differential diagnosis of hypoxaemia is wide and this results in low sensitivity (66.7%) and specificity (60.7%). The oxygen saturation (SaO₂) derived from pulse oximetry has been investigated as a potential test in view of simplicity and patient comfort, particularly for recurrent evaluations. The bias and precision of pulse oximeters has been found to compare favourably with conventional arterial blood samples for oxygen saturations of 80% and above¹¹³. Data on reproducibility in normal subjects suggest a change of 1% in oxygen saturation above 90% and 2% in oxygen saturation below 90%, using pulse and conventional oximeters, indicates

clinical change has occurred¹¹³. This has been further validated in a PAVM population by an institution using Biomedex oximeters, taking the mean of 10 values obtained at one minute intervals¹¹⁴. In these authors' hands, SaO₂ in the erect posture of $\leq 96\%$ gave a sensitivity of 73% and specificity of 35%, and $\leq 97\%$ gave a sensitivity of 100% and specificity of 10%, for detection of PAVMs as defined by selective pulmonary angiography. Their erect SaO₂ was a better predictor of disease than supine saturations. This reflects the preponderance of PAVMs in the lung bases leading to increased shunting in the erect posture due to gravitational pooling of blood¹¹⁵. This study also addressed whether the fall in oxygen saturation on assuming the upright posture, orthodeoxia, could be useful. Although this was no better a predictor of PAVM disease than erect SaO₂ alone, the study demonstrated that the supine-erect SaO₂ was significantly greater in the PAVM rather than non-PAVM group: 1.7% (SD 2.7) compared to 0.3% (SD 1.5) $p=0.01$ ¹¹⁴. Erect oxygen saturation and orthodeoxia measured by pulse oximetry are therefore published methods of identifying individuals requiring further investigation for possible PAVMs but are not complete diagnostic investigations, due to insufficient sensitivity and specificity⁸⁶.

1.2.2.2. CHEST X-RAY

Chest X-ray (CXR) has also been suggested as a screening investigation since 40-83% of PAVMs are revealed by CXR when interpreted by experienced radiologists with a high index of clinical suspicion^{86 98 101 111 116} (see Figure 1.4). However there is a lack of specificity^{100 101} and insufficient sensitivity, even when lateral CXR is used in addition to posteroanterior (PA) CXR, particularly for small or

diffuse PAVMs^{101 116}. It is reported to be better than oxygen saturation for detecting PAVMs which are large enough to embolise⁸⁶. It is important to determine therefore the clinical significance of PAVMs smaller than this (see Section 5.1.2).

1.2.2.3. THORACIC CT SCANNING

Thoracic CT scans can detect small as well as large lesions and identifies thrombosed and recanalised structures. Exposure to radiation and a lack of specificity are problems^{112 117} (see Figure 1.5). Helical scans with 1mm slices at 5mm intervals had 98.2% sensitivity, compared to pulmonary angiography, for detecting PAVMs in one series¹¹².

1.2.2.4. INTRAPULMONARY SHUNT INVESTIGATIONS

Intrapulmonary shunts can be detected using the 100% inspired oxygen breathing method, radionuclide scanning using technetium 99m labelled albumin macroaggregates or microspheres, or by contrast echocardiography.

Some have found that results obtained with the 100% oxygen breathing method and radionuclide scanning are similar¹¹⁸. Others however have found that 100% oxygen breathing tends to underestimate the shunt probably as the “driving pressure” for oxygen diffusion is increased due to a much higher partial pressure of oxygen in the alveoli when breathing 100% oxygen rather than room air¹¹⁹. Recently radionuclide shunting of more than 3.5% has been found to have 87% sensitivity and 61% specificity for detecting PAVMs compared to pulmonary angiography, the “gold standard”¹¹⁴. Comparable figures, allowing for differences in patient selection,

exist for 100% oxygen breathing⁸⁶. The debate about which method is better therefore continues.

Contrast echocardiography utilises the fact that with an intracardiac shunt microbubbles are seen almost immediately in the left ventricle whereas with an intrapulmonary right to left shunt there is a 2-5 second delay¹²⁰. It has been found to be suggestive of intrapulmonary shunt in patients who have no PAVMs on pulmonary angiography- in 40% of patients in one series¹¹¹. It is not clear yet what the clinical relevance is of shunts detected by contrast echocardiography but not seen at pulmonary angiography⁹⁵.

1.2.2.5. PULMONARY ANGIOGRAPHY

This is still the gold standard investigation and is used as part of the therapeutic embolisation procedure (see Figure 1.6). It should be performed in all patients in whom clinical data and CT scanning or shunt measurements are highly suggestive of the presence of PAVMs⁹⁵. Despite pulmonary angiography being the gold standard, it can fail to detect PAVMs identified by CT scanning and shunt investigations, particularly contrast echocardiography^{86 111 112}. It is thought that this is due to these lesions being too small for detection¹²¹ since the abnormal shunts persist after apparently successful embolisation in about 60-75% of PAVM patients^{114 122}. The clinical relevance of these lesions is unknown (see Section 1.2.3.1).

1.2.3. TREATMENT

1.2.3.1. EMBOLISATION

Embolisation has now replaced surgery as the mainstay of treatment as it allows parenchymal sparing, which is particularly important as the disease is often recurrent. There is also a lower procedural morbidity. It is performed at the time of pulmonary angiography by positioning either a detachable balloon or coil in the vessel leading to the PAVM (see Figure 1.6)⁹⁵. PAVMs with a feeding vessel of 3 mm or more in diameter are amenable to embolisation¹²¹. Thrombus then forms around the device. In the majority of patients it is efficacious, in terms of shunt measurement, although is operator dependent, requiring a highly skilled radiologist⁹⁵. Recurrence of PAVMs occurs either due to recanalisation of original PAVMs through poorly occluded vessels, or due to new pulmonary or systemic feeding vessels supplying the original PAVM¹⁰⁷, or because of new PAVMs forming due to the recurrent nature of the disease, particularly likely with the removal of a low resistance shunt⁹⁵. Patients therefore require follow up after a few months and often several treatments are needed^{98 122 123}.

Both balloons and coils must be small enough so that they can be placed sufficiently distally, thereby avoiding occlusion of a vessel leading to a normal capillary, otherwise pleurisy due to pulmonary infarction would occur. It is argued that balloons are less likely to cause this⁹⁷ but around 10% of patients were found to develop pleurisy in 3 series, 2 using coils and 1 using balloons^{97 107 122}. They must not be too small however otherwise they may embolise through the PAVM

themselves^{107 122 123}. Other complications that may arise include paradoxical embolism, angina and cardiac arrhythmias due to air embolism⁹⁷, and postembolisation haemothorax⁹⁸. After treatment of a PAVM, pulmonary hypertension and right heart failure due to high cardiac output from a coexistent hepatic AVM may be unmasked by removal of the low resistance shunt^{107 123}.

Embolisation of vessels less than 3 mm in diameter has been attempted with more difficulty⁹⁵. In one study patients with diffuse PAVMs showed no significant change in hypoxia after embolisation of large PAVMs and 5 episodes of CVA/TIA occurred in the 12 treated patients¹²⁴. This, along with data presented in this thesis (see Section 4.1.1.2), suggests that microscopic PAVMs are clinically relevant, but further data is required.

1.2.3.2. SURGERY

PAVMs occur subpleurally in more than half of patients and so are amenable to video-assisted thoracoscopic wedge excision of the PAVM. This reduces the complications of surgery¹²⁵. Bilateral lung transplantation for diffuse PAVMs in a patient without HHT has been described as being successful in that ABGs and a 6 minute walk test were normal 8 months after transplantation¹²⁶. However long term follow up in this case is not known. In patients with diffuse disease it is suggested that the 2 year survival rate is poorer in those treated by lung transplantation than those treated conservatively¹²⁴.

1.2.3.3. ANTIBIOTICS

Prophylactic antibiotic therapy prior to undergoing procedures that can evoke bacteraemia is recommended in all those suspected of having PAVMs¹²⁷. In view of concern over the significance of microscopic PAVMs this is being extended to include those in whom PAVMs have been treated unless they are one of the few in whom no residual shunts remain⁹⁵.

1.2.4. PREGNANCY AND OESTROGEN

The majority of women who are going to develop HHT have done so by the time they reach child-bearing age³⁵. Pregnancy tends to exacerbate HHT manifestations, in particular pulmonary haemorrhage and complications due to shunting through PAVMs¹⁰⁶, but deterioration is also seen in hepatic, cerebral, and mucocutaneous vascular beds^{106 128}. Various mechanisms are postulated for the deterioration. The increased cardiac output and vascular volume¹²⁹ may precipitate haemorrhage from abnormal vasculature¹⁰⁶. An alternative possibility is that the associated fall in systemic and pulmonary vascular resistance¹²⁹ may accelerate the dilation of existing lesions or precipitate their initial production¹⁰⁶. The latter mechanism would account for the majority of the complications seen.

In addition to the change in haemodynamic state during pregnancy potentially altering the course of the disease, there may also be a localised hormonal influence on the vessels. Oestrogen secretion increases markedly during pregnancy such that towards the end of pregnancy the equivalent of 20mg oestradiol/day is produced¹³⁰. Very high levels and very low levels of oestrogen have been shown to worsen HHT.

Very low levels (<50micrograms of oestradiol equivalent per day) are thought to account for worsening of epistaxis in HHT patients during the low oestrogen states of the 5 days in the menstrual cycle prior to menstruation, the menopause^{48 131}, after oovarectomy¹³¹, and on the low dose oestrogen oral contraceptive pill¹³². Very high levels of oestrogens (20mgs/day) are postulated as accounting for the increased frequency of PAVMs in females¹⁰⁶, and possibly the worsening of PAVMs during pregnancy¹⁰⁶. In spite of these findings, there is evidence that an intermediate dose of oestrogen at a “high contraceptive dose” of 50micrograms oestradiol/day is definitely beneficial for HHT associated GI bleeding⁶⁹ and doses of up to 5mg oestradiol/day may be efficacious for HHT associated epistaxis^{43 48 56}.

The mechanisms by which oestrogens have these effects are likely to vary according to dose. The benefits of high dose oestrogens in management of epistaxis has been ascribed to alterations in the vascular wall by induction of squamous metaplasia (see Section 1.1.4.1). This mechanism could not be responsible for the improvement in gastrointestinal bleeding at lower doses. Possible alternative explanations might include upregulation of transcription from the remaining *endoglin* gene in endoglin-deficient HHT patients (three oestrogen response elements have been identified in the *endoglin* promoter¹³³), a shift towards a prothrombotic tendency¹³⁴, or indirect effects on vascular function.

1.2.4.1. IMPLICATIONS FOR MANAGEMENT OF HHT PATIENTS

The current recommendations arising from these considerations are two-fold. First, the van Cutsem⁶⁹ dose of female hormones (50microgram ethinyloestradiol, 1mg

progesterone) is used to treat transfusion-dependent patients with gastrointestinal bleeding. More relevant to this study is the recommendation that, recognising the potential risk of pregnancy, whether or not it is due to the high oestrogen state, women should be screened for PAVMs ideally prior to, and, if necessary, during pregnancy. Physiological changes in pregnancy may mask hypoxaemia due to PAVMs (an elevation in the mixed venous oxygen content results from the elevated cardiac output and progesterone stimulation of the respiratory centre) and rescreening soon after delivery is warranted⁹⁵. Embolisation is safe up to and including the third trimester¹⁰⁵.

1.3. GENETICS OF HHT

Knowledge of the genetics of HHT is enabling understanding not only of the mechanisms underlying HHT but also, more broadly, those involved in vascular development and homeostasis generally, with far reaching clinical implications in many disease processes. It also establishes whether or not there is any correlation between HHT genotype and phenotype which would have important implications for clinical screening and prognosis.

HHT is known to result from mutations in two genes. An HHT locus, *HHT1*, was assigned by linkage studies to chromosome 9^{35 135}. The disease gene was later identified as *endoglin*¹³⁶. The initial chromosome 9 linkage studies indicated that a further gene existed³⁵ and later a second HHT locus, *HHT2*, was mapped to chromosome 12^{137 138}. The mutated gene was ascertained to be *ALK1*¹³⁹. The existence of a third gene was suggested by linkage studies on one HHT family with extensive hepatic but no pulmonary involvement. This pedigree was unlinked to both *endoglin* and *ALK1*. The locus and gene have not been identified to date¹⁴⁰.

1.3.1. PHENOTYPE-GENOTYPE CORRELATIONS

Could there be any correlation between phenotype and genotype? Several reports suggested that PAVMs only occurred in *endoglin*-linked families and not in those unlinked to *endoglin*^{141 142}. The implication of this would be that only *endoglin* families need to be screened for PAVMs. However although PAVMs are significantly more common in *endoglin*-linked families^{82 143 144} they do not occur exclusively in them^{35 37}. The family found to be unlinked to *endoglin* and *ALK1*

had no pulmonary involvement¹⁴⁰. However this is evidence from only one family and, as will be shown, my data shows otherwise. Moreover affected individuals in the same family can have very differing severities of disease so it is difficult to make a prognosis about the course of disease in any individual^{36 37}.

1.3.2. ENDOGLIN

The endoglin gene lies on the long arm of chromosome 9^{35 135}. Various mutations including deletions and insertions, missense mutations and point mutations generating premature stop codons, have been identified^{37 136 145-148}. Almost all mutations have been found to be unique to a particular family^{37 136 145-148} making mutational screening programmes problematic. In some *endoglin*-linked families no mutations have been found, studying all coding regions and splice sites of the gene³⁷. With the apparent clustering of premature termination mutations, it was hypothesised that a truncated protein is translated which disrupts normal endoglin function via a dominant-negative effect¹⁴⁵. However the identification of missense, frameshift, and nonsense mutations which all lead to very low or undetectable levels of messenger RNA transcripts suggested the prevailing nature of the mutations is to create a null (nonfunctional) allele^{37 148}— the haploinsufficiency model. Furthermore when mutant proteins are translated they only exist transiently within the cell, not reaching the cell surface¹⁴⁶. The haploinsufficiency model is now widely accepted as the mechanism by which these mutations act¹⁴⁹.

This haploinsufficiency model has been further confirmed by a recent murine model of HHT¹⁵⁰ where manifestations of HHT were seen in some mice expressing a

single copy of the endoglin gene. Previously embryonic lethality at gestational day 11.5 in mice completely lacking in endoglin, i.e. homozygous endoglin null mice, of 129/SvJ embryonic stem (ES) cell origin, was shown to occur due to defective vascular development¹⁵¹. Lethality in endoglin null mice was also shown in ref.¹⁵⁰ again due to defects in angiogenesis and also abnormalities in heart development. In these mice (129/Ola ES mice bred into C57BL/6 mice), however, death occurred at gestational age 10-10.5 days suggesting that phenotypic differences between endoglin null mice of different genetic backgrounds may be due to modifier genes and/or epigenetic factors. The role of modifier genes and epigenetic factors was further suggested when observing the heterozygous mice¹⁵⁰ where the variability of organs affected, age of onset, and severity of disease was noted both between mice of different genetic backgrounds (129/Ola mice were most severely affected) and between mice of the same genetic background. This parallels human HHT and is thus an elegant and potentially extremely useful HHT model.

Endoglin (CD105) is predominantly expressed on all types of vascular endothelium and is recognised as an endothelial cell marker¹⁵². It is also expressed on the syncytiotrophoblast (responsible for hormone synthesis and nutrient exchange) of the placenta¹⁵³, fibroblasts^{154 155}, smooth muscle cells¹⁵⁶, transiently on endocardial cushion tissue mesenchyme during heart septation and valve formation^{150 157}, pre-erythroblastic bonemarrow mononuclear cells¹⁵⁸, activated macrophages¹⁵⁹ (it is upregulated on these compared to peripheral monocytes)¹⁵⁹. It is upregulated on endothelial cells in tumour vasculature¹⁶⁰, inflamed tissues including skin lesions

e.g. psoriasis^{160 161}, in response to ionising irradiation¹⁶², and in fibroblasts in pulmonary fibrosis¹⁶³.

The haploinsufficiency model allows screening for possible *HHT1* patients to be performed by measuring the level of endoglin expressed on peripheral blood activated monocytes and umbilical vein endothelial cells. In the U.K., in contrast to the United States and Canada, ethical approval has not been granted for screening of newborns by measuring endoglin expressed on umbilical vein endothelial cells. *HHT1* patients have decreased endoglin levels whereas *HHT2* patients and unaffected individuals have normal levels of endoglin⁹⁵.

It had been suggested that *endoglin* mutations were loss-of-function alleles requiring somatic inactivation of the normal allele for vascular lesion development, in other words a two hit phenomenon similar to that hypothesised for tumour suppressor genes¹³⁶. However reduced levels of endoglin are also seen on the endothelial cells of apparently normal vessels of *HHT1* patients compared to the levels seen in unaffected individuals and also compared to the normal levels of endothelial cell marker, PECAM1, in HHT patients¹⁶⁴. The same levels are also seen in abnormal vessels making it unlikely that the theory is correct that AVMs occur when the normal copy of the endoglin gene is inactivated by a second hit¹⁶⁴.

1.3.3. ENDOGLIN PROMOTER

The promoter region of human *endoglin* was recently cloned and characterised¹³³. The relevant elements that participate in *endoglin* transcription were determined by comparing the activities of a series of *endoglin* promoter fragments. The fragments

were inserted into a reporter vector containing the *luciferase* gene. BAEC (bovine aortic endothelial cells) were transiently transfected with the constructs and luciferase activity measured. The largest construct, -851/+350, i.e. spanning from 851 base pairs upstream of one of the major transcription start sites to 350 base pairs downstream of it, had similar activities to its -400/+341 and -224/+341 derivatives (about 700 fold increase over a promoterless vector). However the construct -851/-400 showed no promoter activity and plasmids -141/+341 and -81/+350 showed a decreased level of activity compared to -400/+341 and -224/+341. This suggests that the region 81 base pairs upstream of the transcription start site can act as a basal promoter with additional regulatory elements in the 319 base pair fragment immediately 5' of this required for optimal transcription. An Alu sequence, often located upstream of regular transcription units, was found at -927/-666 supporting the view that the major transcription regulatory element was downstream of this.

Many transcriptional start sites have been identified within the -400/+341 fragment. Unusually no consensus TATA or CAAT boxes have been found but instead guanine and cytosine (GC) rich boxes at -47/-29 and -5/+16 have been found along with an Sp1 binding site at -37. These are a feature of TATA-less genes which have multiple transcription start sites¹⁶⁵. The lack of TATA and CAAT boxes is a common feature in components of the TGF β superfamily¹⁶⁶⁻¹⁷⁰

Several potential transcription factor binding sites, also present in other endothelial gene promoters, have been identified. These include NF κ B, GATA, AP2, Sp1, and ets and all lie within the 741 base pair fragment. Interestingly NF κ B sequences are responsive to mechanical forces, by regulating shear stress response element-

dependent gene expression¹⁷¹, which could account for a potential role in regulating *endoglin* transcription when endothelial cells are subjected to haemodynamic forces of the bloodstream, similar to mechanisms occurring in other endothelial genes¹⁷¹. GATA and ets are seen in other endothelial genes and, more specifically, AP2 and Sp1 sites are present in promoters of other components of the TGF β superfamily (TGF β 1, β 3, biglycan (a small proteoglycan which interacts with TGF β), and TGF β receptor types I and II)^{166 169 170 172-174}. An ets family member also regulates expression of TGF β receptor type II¹⁷⁵.

Other potential TGF β responsive elements were identified within the 741 base pair fragment including 2 sites for the TGF β -related signalling proteins, Smads, 10 TGF β activation elements, 6 TGF β control elements, and 2 TGF β inhibitory elements. In keeping with this, promoter activity was found to be stimulated by TGF β 1¹³³.

Interestingly 3 oestrogen response elements, 1 glucocorticoid response element, and 5 vitamin D response elements were found in this fragment. This could explain, at least partially, the contrasting effects of low, medium, and high oestrogen states, as discussed in Section 1.2.4.

It has recently been shown that two of the Smads, Smad3 and Smad4, components of the downstream signalling pathway for the TGF β superfamily, bind to specific DNA sequences inducing transcription of downstream target genes. They mimic the transcriptional effect of TGF β 1 on the *plasminogen activator inhibitor 1* (*PAI-1*) promoter (the best characterised TGF β 1-inducible promoter). Although specific Smad-response elements are present in the *PAI-1* promoter, interestingly in at least

one study, mutations which eliminated particular Smad DNA binding sites did not interfere with TGF β 1-dependent or Smad-dependent transcriptional activity. However both TGF β 1-dependent and Smad-dependent transcription was eliminated by mutation of adjacent AP1 sites¹⁷⁶. It would appear from this that Smads alter transcription through sequence-specific DNA binding and through potentiating AP1-dependent transcription¹⁷⁶. A Smad4 mutation identified in a human pancreatic carcinoma abolished Smad4 DNA binding activity, preventing transactivation of TGF β responsive genes. This emphasises the importance of the TGF β superfamily and Smads in carcinogenesis¹⁷⁷. The finding of two Smad responsive elements in the *endoglin* promoter highlights the role of endoglin in TGF β superfamily signalling.

1.3.4. *ALK1*

The activin receptor-like kinase (ALK1) gene lies in the pericentromeric region on the long arm of chromosome 12^{137 138}. Mutations in this gene, like in *endoglin*, appear to be unique to a particular family. They include missense and nonsense mutations, and deletions and insertions, and are found in sequences encoding the highly conserved kinase, transmembrane, and extracellular ligand binding domains^{139 178 179}. Some mutant alleles lead to undetectable or low levels of an unstable messenger RNA transcript suggesting that haploinsufficiency is occurring here as in *endoglin*¹⁷⁸⁻¹⁸⁰.

ALK1 is found, like endoglin, predominantly on human endothelial cells and in lung and placenta^{179 181 182}. In the rat it has also been found on splenic macrophages

and in the bone marrow¹⁸³. The distribution of ALK1 is similar to that of endoglin and also TGF β 1 suggesting a close functional association.

As for *endoglin*, embryonic lethality has been demonstrated in homozygous ALK1 null mice occurring at gestational day 10.5-11.5 due, as in endoglin null mice, to vascular development abnormalities¹⁸⁰. However, unlike in endoglin null mice, the mice in ref.¹⁸⁰ which were heterozygous for the *ALK1* mutation showed no evidence of the HHT phenotype. It is not clear whether this reflects intrinsic differences with requirements for human and murine ALK1, or absence of appropriate environmental triggers in the ALK1 null mice.

1.4. TGF β : THE ROLE OF ENDOGLIN AND ALK1 IN THE SUPERFAMILY

The structurally related peptides forming the transforming growth factor β (TGF β) superfamily are factors which play diverse and fundamental roles in growth and differentiation¹⁸⁴. They include the TGF β isoforms, β 1, β 2, β 3, and β 5, the activins, and the bone morphogenetic proteins (BMPs)¹⁸⁴. They are involved in critical vascular processes such as development^{185 186}, angiogenesis¹⁸⁷, and vascular repair¹⁸⁸, and also in cellular proliferation (*endoglin* can modify TGF β 1 inhibition of cellular proliferation via downregulation of c-myc mRNA)¹⁸⁹, stimulation of cell-matrix adhesion¹⁸⁹, haematopoiesis, immune regulation¹⁹⁰, inhibition of oncogenesis¹⁸⁴, and stimulation of wound repair¹⁹¹. TGF β 1 null mice behave very similarly to endoglin null mice and ALK1 null mice dying at gestational age 9.5-11.5 days due to defects in yolk sac vasculogenesis and haematopoiesis¹⁸⁵. Abnormalities in signalling by the TGF β superfamily are thought therefore to contribute to many different disease processes including importantly oncogenesis, atherosclerosis, pulmonary hypertension and fibrosis. Signalling is mediated by two of a series of serine/threonine kinase receptors occurring in two forms, type I and type II¹⁹², in a complex. The type I receptors transduce signals to downstream targets, Smads^{180 193 194}. These receptors however require first to be activated by a type II receptor which itself has been activated by binding of an appropriate ligand^{195 196}. The TGF β isoforms and activin type I bind ligands only on coexpression of type II receptors^{195 197 198}. The ligand specificity of the type I

receptor is defined by the type II receptor¹⁹⁷. For example, for TGF β 1 the type II receptor, T β RII, confers ligand specificity on the type I receptor T β RI. In the BMPs, the type I receptor can bind ligand independently but still requires a type II receptor for transduction of the signal¹⁹⁹⁻²⁰¹. TGF β isoforms and activins can signal through the same downstream components, Smad2 and 3²⁰²⁻²⁰⁴. BMPs activate Smad1, 5, and 8 signalling pathways^{194 205-207}. The possibility that Smad2 has a tumour suppresser function has been raised as *Smad2* lies within a 3 megabase region of *DPC4* on chromosome 18q21.1, a region often deleted in human cancers²⁰⁴.

Endoglin is a trans-membrane glycoprotein ²⁰⁸. It was identified as a component of the TGF β receptor system binding TGF β 1 and β 3²⁰⁹ which was prompted by observations that endoglin had a similar structure and distribution to betaglycan, a type III receptor²¹⁰. Betaglycan can bind TGF β 1, 2, and 3²¹¹. It presents the bound TGF β ligand to the type II receptor which is able to bind the ligand with much greater affinity than it can bind unbound TGF β ²¹¹. This is especially so for TGF β 2 which has a particularly low affinity for the type I and II receptors in the absence of betaglycan²¹¹. Betaglycan is then displaced from the complex by the type I receptor²¹¹. Membrane bound betaglycan thus upregulates cell responsiveness to TGF β unlike soluble betaglycan which binds TGF β but, as it does not enhance binding to membrane receptors, acts as a downregulator of TGF β effects²¹². It had been suggested that endoglin functions similarly by affecting the binding of TGF β 1 and β 3 to the signalling receptors^{209 213 214}. However recently it has been shown

that endoglin in fact needs coexpression of a ligand binding receptor in order to interact with a ligand²⁰⁷. Furthermore it does not alter overall binding to the kinase receptors but mimics the specificity of the receptor with which it interacts²⁰⁷. Since it cannot bind ligand it is not a true receptor but is better defined as an accessory protein for the TGF β superfamily²⁰⁷. Endoglin can bind TGF β 1 and β 3 by interacting with the T β RII receptor and can bind activin-A and BMP-7 by association with activin type II receptors, regardless of which type I receptor is expressed²⁰⁷. It also interacts with type I receptors ALK3 and ALK6 which bind BMP-2²⁰⁷. Endoglin cannot however interact with BMP-7 bound to BMP type II receptors indicating specificity of its action²⁰⁷. Since endoglin does not alter overall binding to kinase receptors and also as it interacts with many different signal transducing type I receptors, it has been postulated that endoglin could function to alter receptor activation of downstream signalling events such as Smad signalling or even could recruit new proteins into the signalling complexes²⁰⁷. This suggests that it could perform a variety of regulatory roles in diverse cell types according to the kinase receptors, ligands, and Smad mediators present²⁰⁷.

ALK1 is a serine/threonine kinase type I receptor that *in vitro* can associate with the T β RII receptor or activin type II receptor, binding TGF β 1 or activin respectively, although with low affinity^{180-182 198}. However ALK1 was defined as an orphan receptor since this receptor association had not been demonstrated *in vivo* and its *in vivo* ligand was not known⁹⁵. ALK1 has now been shown to bind TGF β 1 in conjunction with T β RII receptors in human umbilical vein endothelial cells¹⁸⁰.

ALK1 mediated signalling has also been shown to inhibit TGF β 1 activity mediated by the type I receptor, ALK5, suggesting that the balance between ALK1 and 5 may have a crucial role in angiogenesis¹⁸⁰. It has been shown to activate SmadI and Smad5 pathways^{180 194}, involved with signalling for bone morphogenetic proteins which are implicated in various aspects of development^{205 206}.

1.5. LINKAGE ANALYSIS

Primary references for sections 1.5.1-1.5.4 may be found in references²¹⁵ and ²¹⁶.

1.5.1. ALLELES AND COINHERITANCE

Linkage studies are based on analysing the inheritance of alleles. Alleles are different states of a locus (a discrete area of the genome for which a gene is one example). If two loci lie very close together on the same chromosome they will be inherited together from parent to child and so the child will have the same pattern of alleles as the parent. As the distance between the two loci increases, the chances of them being inherited together decreases, as there is an increasing chance of separation due to crossover at meiosis. This produces a non-parental combination known as a recombinant.

1.5.2. POLYMORPHIC MARKERS

These are published loci of known location in which a number of alleles are present. The type commonly used in linkage analysis, and used in my work, are loci which include a varying number of short tandem repeat sequences with repeats of 2 to 5 base pairs. Because they can be amplified by PCR, they allow more rapid analysis than previously used markers (restriction fragment length polymorphisms). They also have more than 2 different alleles so are usually more polymorphic. If the parents have 4 different alleles, i.e. both are heterozygous, then the allelic inheritance by the offspring is much more informative than otherwise i.e. it is possible to tell exactly which allele the offspring has inherited from which parent. The heterozygosity index (H.I.) provides a direct estimation of how many individuals will be heterozygotes. The polymorphism information content (P.I.C.) estimates the

probability that an offspring's genotype will allow the deduction of which of the affected parent's alleles it has received. These values are published with the marker. It is obviously prudent therefore to choose markers with a high H.I. and P.I.C. in order to obtain informative results. Since this work was initiated, a new technique for large-scale association genetics studies has been developed. This uses single nucleotide polymorphisms (SNPs) which occur once every 100-300 bases throughout the genome. A database of SNPs is available at The National Center for Biotechnology Information (NCBI), website <http://www.ncbi.nlm.nih.gov/>.

1.5.3. GENETIC DISTANCES

In one Morgan of genetic distance there will be, on average, one crossover per meiosis. Its derivative, the centiMorgan (cM), is approximately equal to a physical distance of one megabase assuming a sex-averaged genetic map of 3,000 cM and 3×10^9 base pairs in the haploid genome. However genetic distances are not exactly the same as physical distances due to double crossovers and some areas being more prone to crossovers than others. 1 cM represents a θ of 0.01. θ is the recombination fraction where

$$\theta = \frac{\text{the number of recombinants}}{\text{the total number of offspring}}$$

Under conditions of free recombination, either because loci are on different chromosomes or because they are widely spaced on the same chromosome, θ is 0.5 (the maximum value of θ). As θ decreases it reflects the appearance of fewer recombinants than would be expected by chance alone suggesting that the two loci lie close to each other.

1.5.4. LOD SCORES

To define the allelic inheritance of polymorphic markers known to lie close to the locus of interest (in this study the HHT disease-causing gene), mathematical qualification of the likelihood of marker positions is derived from the lod score. In the case of disease gene analyses, if the same polymorphic marker allele is always transmitted from affected parent to affected offspring, but not to unaffected offspring, it indicates that the marker is likely to lie close to the disease-causing gene in the pedigree as the mutant allele of the gene would be very unlikely to be separated from the marker allele at meiosis. However if the affected parent's alleles are randomly inherited it suggests that the marker is not close to the disease-causing gene. Linkage analyses mathematically quantify the likelihood of two loci being linked, and therefore in this case a marker gene lying close to a disease-causing gene. The lod score Z represents:

$$\text{Lod score (Z)} = \text{Log}_{10} \frac{\text{likelihood of observations if loci linked, distance } \theta \text{ apart}}{\text{likelihood of observations if free recombination } (\theta=0.5)}$$

By definition at $\theta = 0.5$ (free recombination) the lod score is 0. A positive lod score implies that the observation is more likely to reflect linkage than chance with a score of +3.6 or more indicating linkage to a new autosomal location at a significance level of $p \leq 0.05$. A score of only +1.2 or more is required to significantly ($p < 0.05$) prove linkage to a known site. The more negative the lod score, the more likely that the observations are not due to linkage, with a score of -2 or less significantly excluding linkage at a significance level of $p \leq 0.05$. Two point lod scores were calculated in

this work using MLINK and GENEHUNTER computer programmes via the Human Genome Mapping Project, website <http://www.nhgri.nih.gov/HGP/>.

1.5.5. POTENTIAL PITFALLS IN LINKAGE ANALYSIS

Linkage analysis is an extremely powerful tool enabling the identification of the site of a disease-causing gene in families with inherited disease. The strength of the analysis can be reduced by either false generation of inappropriate lod scores, or lack of detection of true linkage. Potential ways in which these problems can result are outlined below, according to whether these derive from errors in family pedigree assignment, disease phenotyping or DNA sample numbering, or whether they reflect inherent limitations of the methodology.

1.5.5.1. MISPHENOTYPING

In the linkage programmes used in the work in this thesis, MLINK and GENEHUNTER computer programmes via the Human Genome Mapping Project, website <http://www.nhgri.nih.gov/HGP/>, affected individuals' results are assumed to reflect inheritance of the disease gene, and therefore the analyses require the definition of affected disease status with 100% accuracy. Early studies of HHT linkage analyses were sometimes hampered or delayed by difficulties in phenotyping individuals, particularly by groups who used epistaxis alone as definite indicator of affected status. Because of this the HHT Foundation International scientific committee reviewed the then criteria for diagnosis of HHT³⁴, and established stricter requirements for a definite diagnosis of HHT²¹⁷. Throughout this work disease status was designated using these stricter international consensus criteria²¹⁷ (see Section 2.1.3). Particular care was taken to avoid misinterpretation of epistaxis and

telangiectasia (see Section 2.1.2.2) with exclusion from further analysis, if necessary, of any individuals where interpretation was difficult (see Section 2.1.2.2). Exclusion of the diagnosis of HHT is more problematic clinically due to age-related penetrance (though, in practice, of less importance in the linkage analyses as the programme parameters may be set to allow for a degree of diagnostic uncertainty in clinically unaffected individuals). At less than 12 years of age, penetrance of HHT is low (less than 80%³⁵), and there is little data derived from analyses if the penetrance value has to be set below 75%. For this reason apparently unaffected children under the age of 12 years were excluded from molecular analysis (see Section 2.1.3). The linkage programme parameters were then used, in this work, to take age-related penetrance into account in older clinically unaffected individuals, setting penetrance values at $p=0.8$ between ages 12 and 35 and $p=0.95$ for individuals aged over 35, according to published data on penetrance in HHT^{30 34 35} (see Section 2.9.2).

1.5.5.2. NON PARENTAGE

Incorrect assigning of parentage as stated by the family, or erroneous sample mixing, is usually quickly highlighted as non-Mendelian patterns of inheritance of marker alleles result. The program MLINK will not proceed to generate lod scores in the presence of non-Mendelian inheritance. In theory Mendelian inheritance could occur in cases of non-parentage if the “false” parent had two alleles of which at least one and possibly both were the same as in the true parent. As over 50 different polymorphic markers with allele frequencies of between 0.1-0.3 have been studied in all of these families in the Shovlin laboratory, always with Mendelian inheritance, the probability of coincidental allele sharing on each occasion of non-parentage is

exceedingly remote (in the region of 0.3^{50} for any one family). The only realistic means in which non-paternity could be an issue is if there was genuine sharing of all alleles by the “false” and “true” parent being identical twins but, in this case, the parents would be genetically identical and linkage analysis would not be affected. Similar considerations apply in the case of offspring, for example adopted children would have no higher chance of sharing their adoptive parents’ alleles than the general population and the probability that we would miss this with over 50 markers studied would again be in the region of 0.3^{50} . We therefore did not consider non-parentage to be an issue in these families studied in this work.

1.5.5.3. INHERENT LIMITATIONS IN LINKAGE STUDIES

The potential of linkage to determine the site of a locus or disease gene critically depends upon the presence of neighbouring polymorphic markers in which informative alleles are faithfully passed on to individuals’ descendants.

1.5.5.3.1. MARKER DISTRIBUTION

Lack of appropriate polymorphic markers was the major limitation to early linkage studies which were only able to compare the location of one gene with a known function to another^{218 219}. The principles of creating a map using anonymous DNA markers throughout the human genome were delineated more recently, but this was initially again restricted to a relatively limited map of widely spaced markers²²⁰. With the discovery of the short tandem repeat polymorphisms, such as (CA)_n repeats, virtually all areas of the genome became accessible to these studies. It is therefore now highly unlikely that lack of a suitable marker would be the reason

why linkage would be missed, even if the locus under study lay in a relative hot spot for recombination.

These considerations are advanced even further with the discovery and utilisation of single nucleotide polymorphisms (SNPs), although these were not used in this study. The human SNP map was published in February 2001²²¹ with 1.42 million SNPs, with a SNP on average every 1,000-2,000 nucleotides^{222 223}.

1.5.5.3.2. REDUCED ALLELE INFORMATIVENESS

If insufficiently polymorphic, there is a significant chance that for any given marker it will be impossible to assess which parent's alleles a particular offspring has received. This is the essential prerequisite for linkage studies, estimated for each marker as the polymorphism information content (PIC)²²⁰. If all markers in the vicinity of a particular locus have a low PIC, particularly in the families derived from common population pools, then the effective distance to the nearest informative polymorphic marker will be reduced. This results in the possibility that linkage might be missed.

Informativeness of markers is also reduced in extended families connected by numerous deceased ancestors, since it is impossible to assess whether the allele apparently segregating with disease is identical to that segregating elsewhere in the pedigree, rather than being brought in incidentally by an unaffected spouse. The latter is more likely to occur if the allele is particularly frequent in the population. Since disease allele frequencies are generally set at 10% to calculate initial lod scores, wherever co-inheritance stems through a deceased ancestor the linkage

programme will assign an artificially high probability of linkage. In this setting lod scores may be misleadingly high. This risk is reduced if the study focuses on large sibships which reflect fewer parental pairs.

1.5.5.3.3. GENOME INSTABILITY

Linkage depends on faithful transmission of alleles from parent to child. This will not occur if the polymorphic alleles undergo a spontaneous mutation during meiosis resulting in apparent non-parentage or, more rarely but of more concern to linkage analyses, apparent inheritance of a disease-associating allele. In general, the spontaneous mutation rate of human genomic DNA is approximately 10^{-9} to 10^{-12} errors per base per generation²²⁴. Short tandem repeat sequences, which were used in this study as polymorphic markers, arose by replication slippage, and might be suspected to be prone to a higher mutation rate than other areas of the genome. However their spontaneous mutation rate was estimated as 0.03% in over 17,000 meioses²²⁵. Certain diseases, such as familial non-polyposis colon carcinoma, generate genome instability causing an increased spontaneous mutation rate, thus reducing the power of linkage analysis. This has not been seen in the tens of thousands of parent-child allele transmissions in HHT performed in the Shovlin group, and was therefore not considered a limitation to linkage analysis in this disease.

1.5.6. LINKAGE STUDIES VERSUS ASSOCIATION STUDIES

Linkage analysis is an excellent tool for studying diseases determined by the action of a single gene. However it is less useful in non-Mendelian diseases, arising from the interaction of a number of genes and environmental influences, as studies of

small numbers of extended pedigrees containing large numbers of affected individuals will not yield the same results given that the mode of inheritance is unknown and a number of genes with segregating alleles in each generation is involved. One tool for tackling this is the association study. In this, genotypes in cases and controls at a genetic marker of interest are compared and the difference in the allele frequencies between cases and controls interpreted as evidence in favour of association of one or other allele at the marker of interest with disease. Unlike a linkage study it does not have to be carried out within a family group. Whereas linkage is a property of loci, separated by the same genetic distances in all populations, association (also known as linkage disequilibrium) is a property of alleles ²²⁶.

The feasibility of analysing more complex genetic traits demands a higher density map of polymorphisms than the series of short tandem repeats sequences used in linkage studies in monogenic disease. With a SNP on average every 1,000-2,000 nucleotides^{224 225}, SNPs are proving particularly useful in association studies looking at diseases with a complex, multigene basis. In addition each SNP can be classified as to whether it is likely to be directly functional as many occur in coding sequences (when the encoded protein may be altered) or regulatory regions of the genome²²³.

Understanding of the genetic basis of disease, whether the disease is due to a single gene disorder, multigenic, or a combination of multigenic and environmental factors, through linkage and association studies should lead to enormous therapeutic

possibilities from gene therapy, through pharmacogenomics, to primary prevention²²⁷.

1.6. INTRODUCTION TO EXPERIMENTAL WORK

From our current knowledge of the molecular mechanisms underlying HHT and its clinical implications, my work has been based on two separate hypotheses. The first is that mutations in the *endoglin* promoter region could account for disease in *endoglin*-linked families. The second separate hypothesis is that pulmonary involvement could occur in HHT families unlinked to *endoglin* and *ALK1* but instead linked to a third gene.

Hypothesis 1: Mutations in the endoglin promoter region could account for disease in endoglin-linked families

Not all mutations have been found in *endoglin*-linked families studying all coding regions and splice sites of the gene³⁷. Given that haploinsufficiency is the prevailing model in *endoglin*-linked families^{37 148}, it seems likely that mutations in the promoter region of the gene could account for disease.

The 8 HHT families identified for this work were chosen as 4 had previously been shown to be linked to *endoglin* but no mutations had been found by Southern blotting, cycle sequencing of all known exons and intron/exon junctions, and cycle sequencing of the entire cloned DNA³⁷. No mutations had been found in the other 4 families, using the same techniques, but they were too small to allow significant linkage to an HHT gene to be made³⁷.

In other work the term “mutation” has been used to describe any permanent variation between template and daughter sequences in DNA polymerisation. However in this

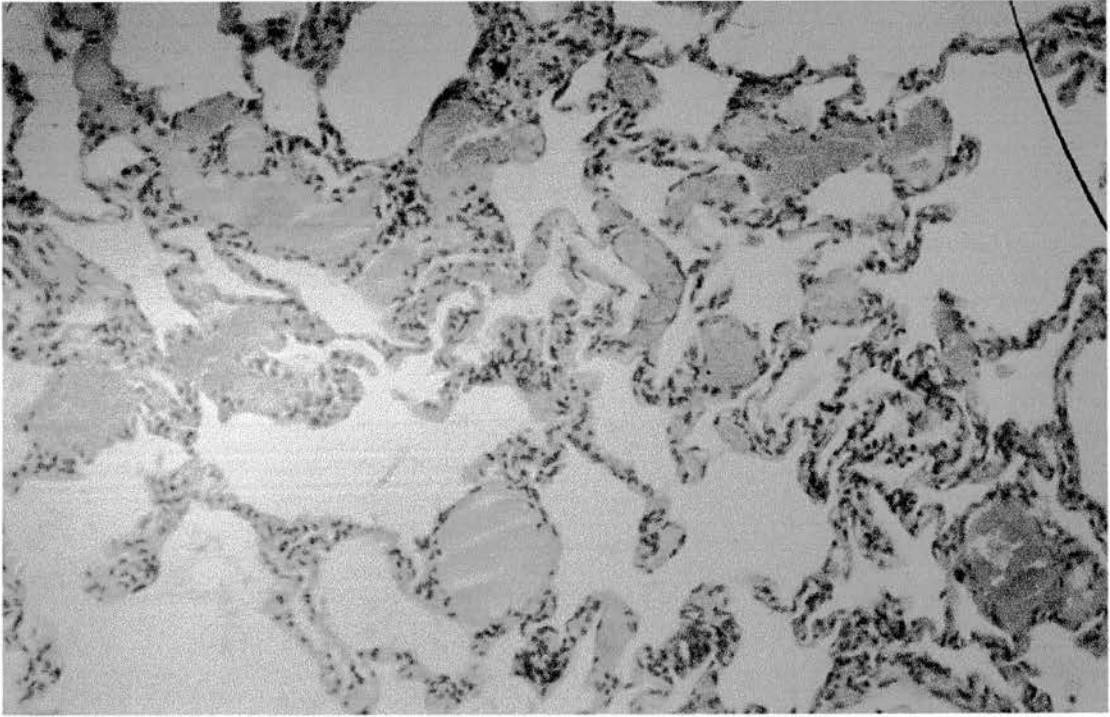
work mutation is defined as this variation occurring in an allele that is only present cosegregating with a particular disease trait and for which there is independent evidence that this variation is the cause of the disease. In this work a variation in an allele which is incidental to the gene product and not a disease causing mutation is described as a polymorphism.

Hypothesis 2: HHT families with pulmonary involvement may be unlinked to the PAVM gene, endoglin, and may be unlinked to both known HHT genes, endoglin and ALK1

The work of Piantanida¹⁴⁰ suggests that a third gene for HHT could exist. While no pulmonary involvement was seen in this pedigree, PAVMs have been shown not to occur exclusively in *endoglin*-linked families^{35 37}. The hypothesis, therefore, seems plausible.

The HHT families chosen for this study were identified as being suitable as they had pulmonary involvement and initial analysis in 1992 suggested that they were unlinked to *endoglin*³⁵. Linkage analysis was undertaken using the principles as discussed in Section 1.5.

FIGURE 1.1 HISTOLOGY OF MICROSCOPIC PAVMS



Haemotoxylin and eosin stained slide, viewed by light microscopy, showing microscopic PAVMs.

I would like to thank Dr D. Lamb and Professor W. MacNee for this figure.

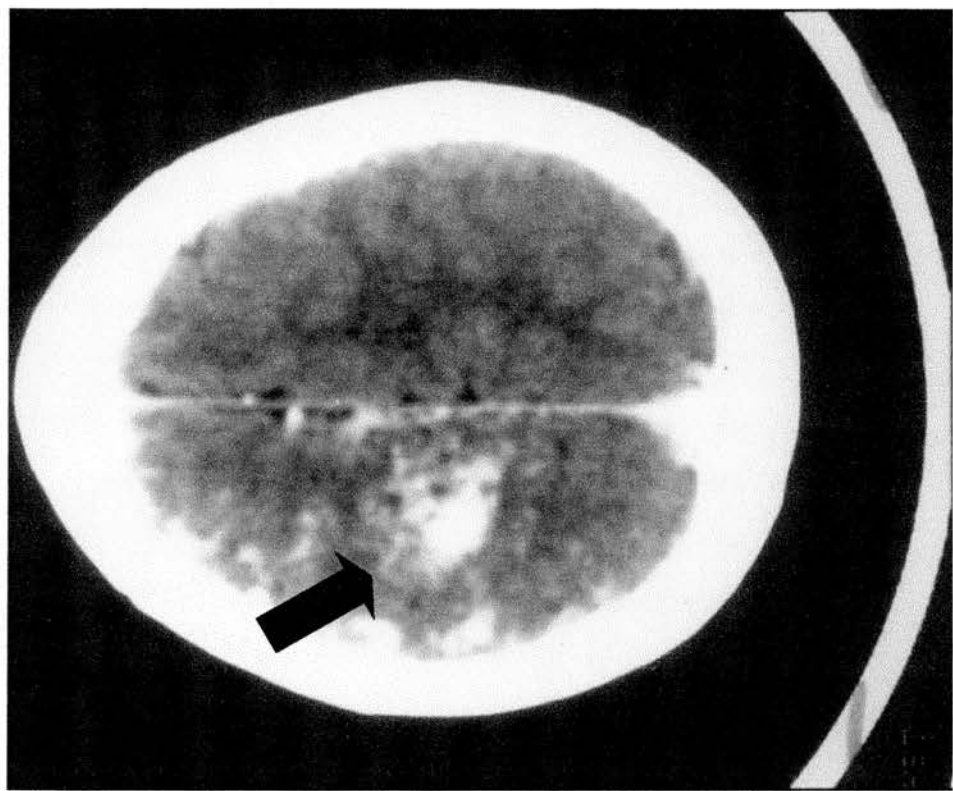
FIGURE 1.2 TELANGIECTASIA OF THE LIPS



Telangiectasia of the lips.

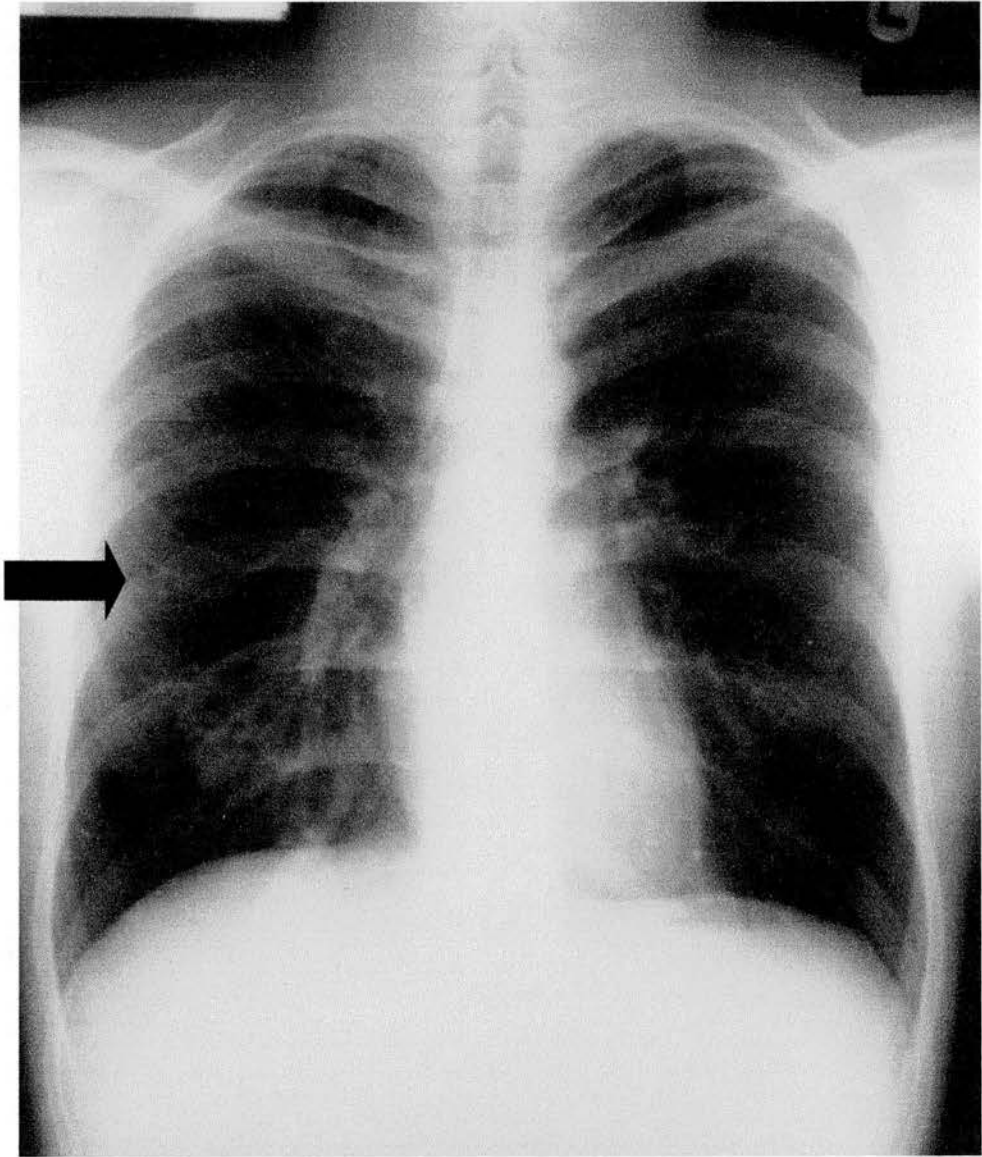
I would like to thank Dr C.L. Shovlin for this figure.

FIGURE 1.3 CT HEAD SCAN SHOWING CAVM



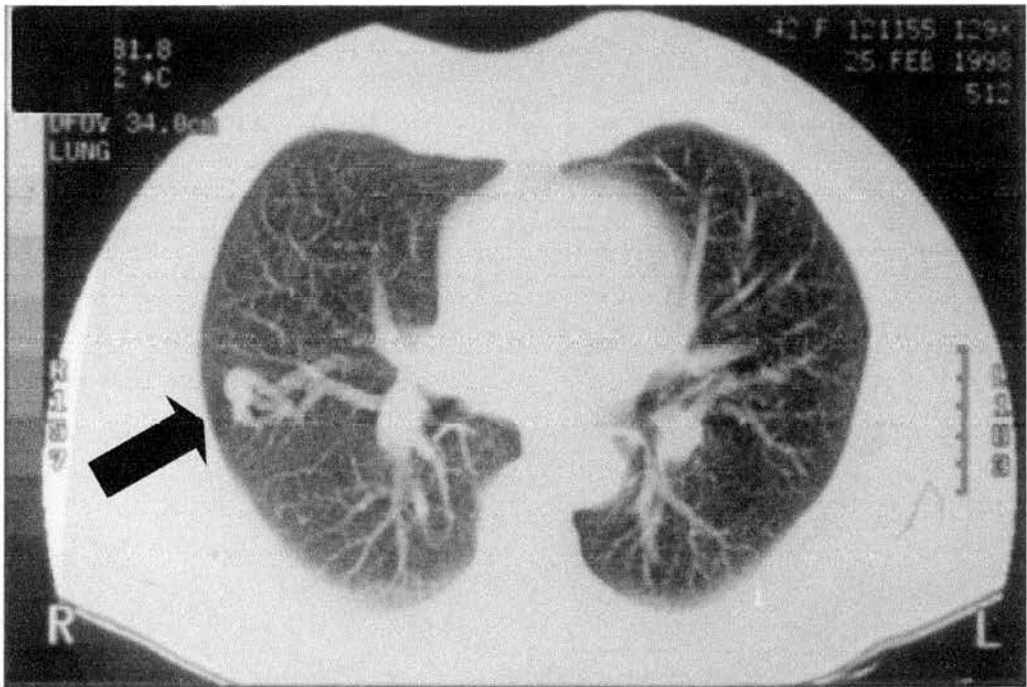
CT head scan showing CAVM marked by the arrow.

FIGURE 1.4 CHEST X-RAY SHOWING PAVM



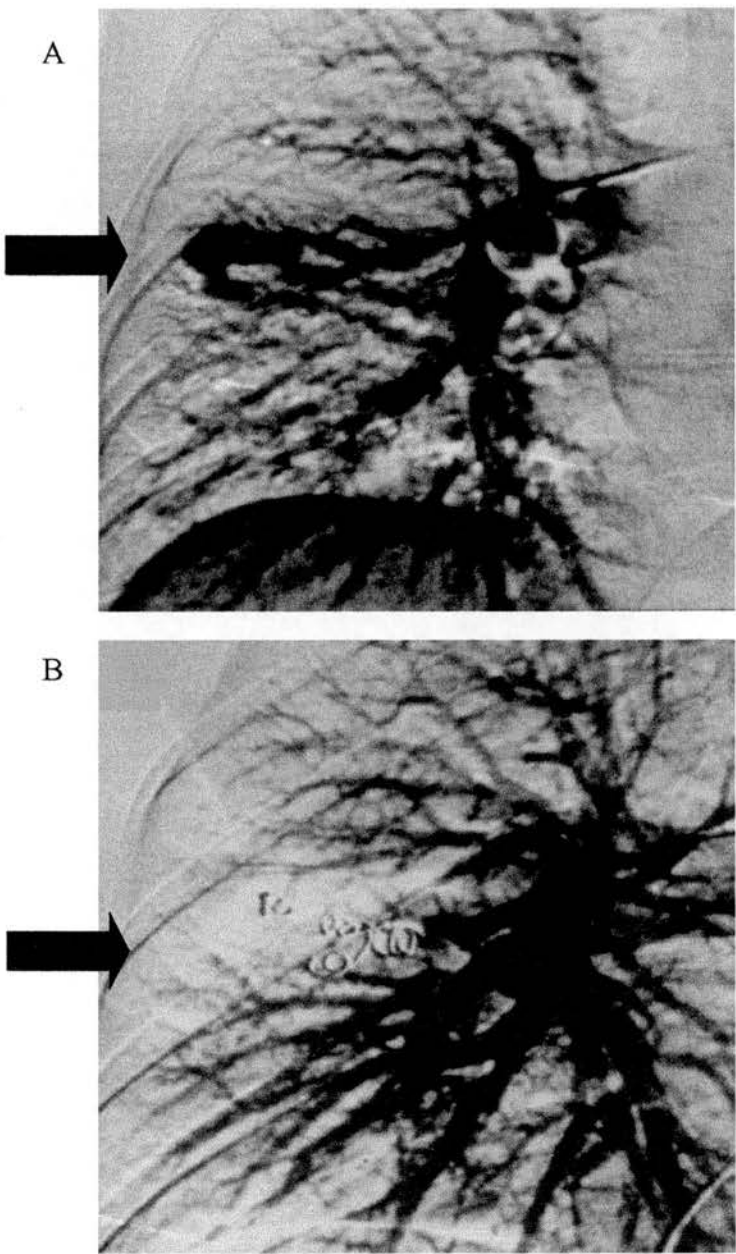
PAVM on chest X-ray indicated by the arrow.

FIGURE 1.5 CHEST CT SCAN SHOWING PAVM



CT scan of chest showing PAVM indicated by the arrow.

**FIGURE 1.6 PULMONARY ANGIOGRAM PRE AND POST
EMBOLISATION**



Pulmonary angiography showing PAVM before treatment in A) and after coil embolisation in B) (marked with arrows). I would like to thank Dr K. McBride for performing the investigation.

2. MATERIALS AND METHODS

Both clinical and molecular techniques were used.

2.1. CLINICAL

2.1.1. ACQUISITION OF SUBJECTS

Current studies were performed with the approval of the Multicentre Research Ethics Committee of Scotland (MREC). Families affected by HHT were initially identified in 1992^{35 37} when one individual from each presented to the Hammersmith Hospital for management of PAVMs. Families of the individuals were clinically evaluated at that time by Dr Claire Shovlin (CLS), my supervisor, in the same way as documented below.

2.1.1.1. LINKAGE ANALYSES

Two of the families (families S & P) were further assessed by myself and CLS six years later. These families were identified as being suitable for further linkage analyses as analysis in 1992 indicated that they were not linked to *endoglin*. Original members were reviewed and additional members were assessed. After linkage results were obtained for these two families, two additional families (families V and T) with PAVMs were studied. They had been reported as being unlinked to *endoglin*^{35 37}. They had been assessed clinically in 1992 by CLS and were not assessed further clinically by myself.

2.1.1.2. PROMOTER ANALYSES

Eight of the families identified in 1992 were studied further. They were chosen for this work as four had previously been shown to be linked to *endoglin* but no mutations had been found. No mutations had been found in the other four families but they were too small to allow significant linkage to an HHT gene to be made³⁷.

These families had been clinically evaluated by CLS in the same way as documented below. They were not assessed further clinically by myself.

2.1.2. CLINICAL EVALUATION

With written consent individuals were evaluated clinically in their own homes with a full medical history. This included questions regarding epistaxis, gastrointestinal bleeding, anaemia, dyspnoea and haemoptysis, and symptoms suggestive of transient ischaemic attacks (TIAs) (see Appendix A). Physical examination for telangiectasia on the lips, face, hands, and on the buccal and nasal mucosae was performed. The presence of central cyanosis and finger clubbing, which could be due to PAVMs, was noted. Full blood counts (FBCs) were performed when appropriate (to assess for anaemia or polycythaemia). If necessary, further information was obtained from GPs and local physicians, with patient consent.

2.1.2.1. SCREENING FOR PAVMS

Screening for PAVMs is recommended for all individuals with HHT⁹⁵ in view of the risks of paradoxical embolism if PAVMs are left untreated and the safe therapeutic methods which exist to limit these complications. As respiratory physicians, in this study, we wanted to perform an assessment in addition to clinical history and examination, in order to direct potentially affected patients to local respiratory services for formal evaluation.

Although physiological considerations suggest that oxygen saturations (SaO₂) detected by pulse oximetry might be too imprecise to use to predict for the presence of PAVMs¹¹³, this method has been previously validated in PAVM patients¹¹⁴.

Thompson *et al*¹¹⁴ used Biomed pulse oximeters and a protocol of measurements for 10 minutes in both the supine and erect postures. In their hands, erect SaO₂ ≤ 96% had a sensitivity of 73% and specificity of 35%, and this value was recommended for a screening cut-off. In addition, in their population, supine-erect SaO₂ was significantly greater in the PAVM rather than non-PAVM group: 1.7 (SD 2.7%) vs 0.3 (SD 1.5%), $p=0.01$.

We elected to base a simple screen on these results. We purchased an Ohmeda 3800 oximeter and used the Thompson protocol of measurements for 10 minutes in both the supine and erect postures. Diagnostic cut-offs for results sufficiently suggestive of PAVMs to advise (with patient consent) referral to local respiratory physicians were established. These were suspicious symptoms (e.g. unexplained haemoptysis) or signs (e.g. clubbing, central cyanosis), SaO₂ erect of ≤ 96% (as recommended by ¹¹⁴), or SaO₂ supine-erect greater than 2SDs beyond the mean in the non-PAVM group in ¹¹⁴, i.e. $0.3\% + [2 \times 1.5\%] = 3.3\%$.

2.1.2.2. ACCURATE PHENOTYPING

Accurate phenotyping is critical. Care was required particularly in evaluating certain manifestations of HHT. It is important that epistaxis is spontaneous and not precipitated by trauma. Two individuals gave a history of nose bleeds which had only occurred after possibly being hit on the nose. They could therefore not be given a definite phenotype and were given the disease status “unknown” (see Section 2.1.3). Telangiectasia may also be misdiagnosed. All telangiectasia were assessed by myself and CLS rather than individual reports of them being accepted. Differentiation from spider telangiectasia can be difficult, although they are

morphologically distinct²²⁸. One individual, with a history of heavy alcohol usage, had to be designated “unknown” (see Section 2.1.3) as her telangiectasia occurred in the presence of such extensive spider telangiectasia that we were not confident to diagnose a subset of HHT telangiectasia with certainty.

2.1.3. DESIGNATION OF DISEASE STATUS

Disease status was designated using the current international consensus criteria²²⁹. Affected status was assigned in the presence of any three criteria: affected first degree relative; recurrent, spontaneous epistaxis; telangiectasia; documented visceral manifestations. If only two of the four criteria were present (affected first degree relative and epistaxis or telangiectasia), unknown status was assigned and the individual excluded from molecular analyses. However if the offspring of the individual was affected, a diagnosis of obligate affected was made. Apparently unaffected children under the age of twelve years were designated unknown, as penetrance is less than 80% at this age^{34 35}.

2.1.4. SAMPLE COLLECTION

With consent each individual had 20mls of blood venesected into a heparinised 50ml Falcon tube. This was for later DNA extraction and, in some cases, leukocyte and plasma preparation. Where appropriate 2.7mls of blood was also placed in an EDTA tube for a full blood count (FBC) to be performed. The samples were collected on trips which lasted several days so were kept at ambient temperatures for about three days. The samples in the Falcon tubes were then placed at -70°C while awaiting DNA extraction. FBC samples were sent to the haematology laboratory at the Royal Infirmary of Edinburgh for analysis.

2.2. DNA EXTRACTION

Standard proteinase K digestion and phenol-chloroform extraction methods were adopted²³⁰ with ethanol precipitation and resuspension in Tris-buffered EDTA:

2.2.1. GENOMIC DNA

Genomic DNA was extracted from all the whole blood samples obtained from families S and P, which had been previously placed at -70°C . Half of the total original sample for all family S members had already been used for leukocyte and plasma extraction. Half of the total remaining sample (half of the total original sample for family P members) was prepared so that more of each sample was available in case of problems. Frozen samples in heparinised Falcon tubes were thawed in a water bath at 37°C . They were diluted in an equal volume of autoclaved phosphate buffered saline (PBS) and centrifuged at room temperature at $3500g$ for 15 minutes. To commence deproteinisation the resulting cellular pellet was resuspended in extraction buffer (see Appendix B) (the volume of this used was three quarters of the original sample volume), and incubated at 37°C for one hour. The buffer contained the ionic detergent 10% sodium dodecyl sulfate (SDS) (Gibco) and EDTA (Sigma) which inhibits DNA endonucleases (DNases). RNaseA (Sigma) was added to the buffer at $20\mu\text{g/ml}$ to remove RNA. After this incubation, deproteinisation was completed with proteinase K (Sigma). Initially it was added to each sample to give a concentration of 20mg/ml and incubated at 50°C for 3 hours swirling regularly. However, despite this being a higher concentration than suggested by Sambrook²³⁰, it was always insufficient, with several more additions of proteinase K usually being required. The samples were incubated after each addition either at 50°C for 3 hours

or 37°C overnight. Double the original volume of proteinase K was therefore added each time which decreased the number of additions required.

RNaseA and other lipid soluble cellular debris was removed by extraction with phenol and/or chloroform. For each extraction an equal volume of the appropriate phenol/chloroform solution was added and mixed by inversion by hand for 10 minutes. This minimised shearing forces on the DNA. The sample was then spun at 2000 rpm at room temperature for 15 minutes and the resultant supernatant removed using cell saver pipette tips, again to decrease shearing forces. The next extraction was then commenced on the supernatant. Two phenol extractions (equilibrated phenol (Sigma)), followed by a phenol:chloroform 50:50 extraction (phenol:chloroform isoamylalcohol 25:24:1 (Sigma)), and finally a chloroform extraction (chloroform isoamylalcohol (Sigma)) were performed.

To precipitate the DNA, 0.1× volume (of the final total volume) of 3M sodium acetate and 2.5× volume (of the volume of supernatant remaining from the final chloroform extraction) of 100% ethanol was added to the supernatant. The DNA precipitated at room temperature and could be lifted out with a sterile glass rod. The precipitated DNA was washed twice in 70% ethanol to remove salts and small organic molecules and air dried. Vacuum drying in a Savant speedvac was tried but tended to make resuspension more difficult. The DNA was resuspended in approximately 500µl of GTE.

The DNA concentration, integrity, purity, and resuspension were assessed by agarose gel electrophoresis of 1µl of extracted DNA in comparison to 1µl of λHindIII

marker. If residual contaminating DNA and protein had been excessive, relevant purification steps could be repeated, although this did not occur. If resuspension was problematic aliquots of the DNA were diluted further with genomic TE, heated to 37°C and vortexed (gently to minimise shearing of the DNA), prior to use for PCR.

2.2.2. PLASMID DNA

Plasmid DNA clones were grown to saturation in 7mls LB (supplemented with 25-50 µg/ml ampicillin for ampicillin selection). After centrifuging at 2500 rpm, at room temperature for 5 minutes, the resulting cellular pellet was resuspended in 150µl Genomic tris EDTA (GTE) (see Appendix B). Bacterial lysis was performed by alkaline lysis²³⁰ by adding 300µl of 1% SDS (to denature all DNA). The reaction was vortexed vigorously and iced. Subsequent addition of 300µl of ice cold 5 molar (M) potassium acetate pH 5.0 followed by vortexing, icing for 5 minutes, and vortexing again, allowed the covalently-closed plasmid DNA to reanneal rapidly, while most of the chromosomal DNA proteins and DNA precipitated with the newly complexed SDS and potassium and was removed by centrifuging for 10 minutes at 4°C. 1µl of 10mg/ml RNaseA, to remove residual RNA, was added to the supernatant and incubated at 37°C for 30 minutes prior to proceeding to one phenol:chloroform 50:50 extraction using the same technique as for genomic DNA (See Section 2.2.1).

Precipitation of the DNA was performed using the same technique as for genomic DNA i.e. ethanol precipitation (See Section 2.2.1). Since plasmid DNA is composed of shorter fragments of DNA than genomic DNA precipitation occurs less readily. It

was enhanced by the addition of 1µl of glycogen to coprecipitate as a carrier molecule and the sample placed at -70°C for 20 minutes followed by centrifugation at 4°C for 20 minutes. The precipitated DNA was washed and dried in the same way as for genomic DNA and resuspended in 50µl of GTE.

The insert DNA was released from plasmid DNA (pGEM-T vector, Promega) by digestion with EcoRI using standard protocols (see Section 2.5.1). The digest was then assessed by agarose gel electrophoresis.

2.2.3. DNA EXTRACTION FROM AN AGAROSE GEL

This was performed using the Wizard PCR Preps DNA Purification System (Promega). PCR products were separated by TAE low melting point agarose gel electrophoresis (see Section 2.4.1.1). The DNA band was cut from the gel using a sterile blade aiming to obtain no more than 300µl of agarose. Exposure of the DNA to UV light was minimised by excising it quickly under a transilluminator UVP with the power in the “lo” position. The sample was melted in a 1.5ml eppendorfer tube by incubation at 70°C. 1ml of Wizard Preps resin was mixed with the melted agarose by pipetting for 20 seconds. Vortexing was avoided to minimise shearing forces on the DNA. The mix was drawn into a Wizard Minicolumn using a syringe. The minicolumn was washed with 1ml of 80% isopropanol and centrifuged at 12000g for 20 seconds to remove residual isopropanol. The DNA was eluted from the minicolumn by applying 50µl of double distilled water (at room temperature if the DNA fragment was less than 3kilobases (kb), at 65°C if the fragment was 3kb or

more), waiting for 1 minute, and centrifuging at 12000g for 20 seconds. The DNA was stored at 4°C.

2.3. PREPARATION OF LEUCOCYTES AND PLASMA

Half of the total volume of whole blood samples from all family S members was used to extract plasma and leukocytes using a Ficoll preparation. Samples were thawed in a water bath at room temperature and diluted with an equal volume of autoclaved phosphate buffered saline (PBS). The same volume as the original blood sample of histopaque 1077 (Ficoll) was overlaid in a 50ml Falcon tube with the diluted blood very slowly to prevent mixing of the two layers. This was centrifuged at 2200 rpm, for 20 minutes at room temperature, with no brake so as not to disrupt the cells. This separated the sample into 4 layers: plasma, leukocytes, histopaque 1077, and erythrocytes from top to bottom respectively. The plasma was removed and stored at -70°C . The leukocytes were removed and washed twice with 50mls PBS, with centrifugation at 1600 rpm for 15 minutes then 1200 rpm for 7 minutes between washes (both times at room temperature with no brake). Leukocytes from this preparation were stored by resuspending the pellets with 1ml RPMI-1640 medium (Sigma) containing previously added foetal calf serum, 10,000 units of a penicillin/streptomycin mix, and L-glutamine (5mls of 200mM stock), and placing on ice. 1ml of foetal calf serum containing dimethylsulphoxide (DMSO) (Sigma), previously iced, was mixed with this and the resulting mixture placed in a cryo tube, iced immediately, and stored at -70°C .

2.4. DNA ANALYSIS

2.4.1. GEL ELECTROPHORESIS

2.4.1.1. AGAROSE

A range of agarose concentrations (0.8-4%) were used according to the DNA fragment size to be studied²³⁰. Low melting temperature products, seaplaque agarose (FMC Bioproducts), were used either if a sample was being gel purified or to enable a 4% gel to be used, as 3% seaplaque agarose, 1% normal agarose (seakem LE agarose (FMC Bioproducts)). Gels were prepared and run in 1× tris borate EDTA (TBE) (Gibco) (see Appendix B) or 1× tris acetate EDTA (TAE) (Gibco) buffers. TAE was always used if gel purification was being performed. 0.5µg/ml ethidium bromide was added to the gel to allow visualisation of the DNA. A loading buffer containing glycerol and bromophenol blue (see Appendix B) was used. Commercially supplied markers (ϕx174HaeIII digest or λHindIII digest) were used. Negative controls were loaded for each reaction.

2.4.1.2. DENATURING POLYACRYLAMIDE

These gels were used to allow size separation of sequencing products, or radioactive PCR products in linkage analysis. Samples were denatured by the addition, for sequencing products, of 5µl of stop solution from the Stratagene Cyclist Taq DNA sequencing kit or 10µl of denaturing gel loading buffer (see Appendix B) for PCR products. Each sample was heated to 95°C for 7 minutes then snap cooled to 0°C. Samples were thereafter kept on ice until loading. 6% gels, used for PCR product separation and some of the sequencing, were made with 42g of urea, 12mls of 10× TBE, 20mls of acrylamide-bisacrylamide 30% (w/v) (37.5:1 solution) and made

up to a final volume of 100mls with double distilled water. For 5% gels, used to resolve larger sequencing products, 10mls of 50% longranger solution (Sigma) was used instead of the acrylamide-bisacrylamide 30%, with more water required, again giving a total volume of 100mls. All gels were polymerised with 50 μ l N,N,N',N'-tetramethylethylenediamine (TEMED) and 500 μ l of freshly made 10% ammonium persulphate. Gel plates were prepared by washing thoroughly with decon neutrocon detergent followed by 100% ethanol, and covering with a silicon containing solution, sigmacote (Sigma) to allow easy pouring and removal of the plates from the gel. Gels were run in 0.6 \times TBE at 60W for one and a half to three and a half hours according to the size of the product being run. After transfer to Whatmann 3MM paper they were dried at 80°C on a vacuum slab dryer (BioRad).

2.4.2. DETECTION SYSTEMS

Agarose gels were visualised by ultraviolet illumination with a UVP transilluminator and detection with Grab-IT computer software. Polyacrylamide gels were exposed to Biomax Kodak film, using an intensifier screen, at -70°C. The film was developed in an X-ograph imaging systems processor. Areas of interest on some of the linkage gels, detected by Kodak film, were then analysed using a Molecular Dynamics phosphorimager to obtain intense banding rapidly.

2.5. ENZYME MANIPULATIONS OF NUCLEIC ACIDS

2.5.1. RESTRICTION ENDONUCLEASE DIGESTS

PCR and cloned products were purified by ethanol precipitation (see Section 2.2) and resuspension in double distilled water to remove buffers and other reagents which could interfere with the restriction enzyme activity. To maximise enzyme activity DNA was digested in a 20-50 μ l volume with a tenth volume of the appropriate 10 \times buffer, and less than a tenth enzyme volume (to prevent glycerol inhibition of the reaction) using a 2-10 fold excess of enzyme over DNA. Acetylated bovine serum albumin (BSA) was added to a final concentration of 0.1mg/ml, as recommended by the manufacturer, to improve the efficiency of digests of impure DNA. The digests were incubated for 1-4 hours at the optimum temperature as specified by each enzyme (for the majority 37°C, for BsrI 65°C). For double digests enzymes with compatible buffers were selected.

2.5.2. LIGATIONS

Tth/vent polymerase based PCR products have blunt ends. These were changed to cohesive ends, to improve the efficiency of the subsequent ligation, by the addition of an adenine nucleotide. This was achieved by incubating a particular PCR product in a Taq based 50 μ l reaction at 75°C for 2 hours, containing a fifth volume 5MM dATP. The ethanol precipitated product was then ligated using DNA ligase from the pGEM-T vector Promega kit to form phosphodiester bonds between adjacent 3'hydroxyl and 5' phosphate groups of cohesive ends. A 10 μ l reaction containing 1 μ l of 10 \times ligase buffer, 1 μ l of pGEM-T vector, 1-3 units of DNA ligase, and the

DNA to be ligated was incubated at 15°C overnight to optimise annealing versus enzyme activity.

2.5.3. LABELLING OF PCR PRODUCTS

PCR products were labelled using Redivue ^{33}P γ -ATP (Amersham) for linkage analyses and thermal cycle sequencing. For linkage analyses only one primer and therefore one PCR product strand was labelled, instead of both, as this is generally known to achieve better resolution of the two alleles. 4 μg of the 5' primer was endlabelled at 37°C for 1 hour in a 25 μl reaction containing tenth volume of 10 \times polynucleotide kinase buffer, 6 μl of Redivue ^{33}P γ -ATP, and 10 units of T4 polynucleotide kinase. A similar reaction was used to label the primer for thermal cycle sequencing except only 100ng of primer was labelled to improve specific activity.

2.6. DNA AMPLIFICATION: POLYMERASE CHAIN REACTION

Reactions were mainly performed in thin walled microtubes (Advanced Biotechnologies) where efficient heat conduction through the tubes allowed the use of short cycle times. They were amplified in a PTC-200 thermal cycler (MJ Research) which has a heated lid preventing evaporation, so no mineral oil was required. In the first few months of my work 500µl eppendorfer tubes were used instead, with amplification in an early Perkin-Elmer machine. Mineral oil and longer cycle times were required. Either enzyme derived from the thermophilic bacteria *Thermus aquaticus* (Taq) or two enzymes also from thermophilic bacteria *Thermus thermophilus* (Tth) and *Thermococcus litoralis* (vent) were used in reactions and are discussed below. The optimum concentration of template DNA was assessed by performing identical PCRs at different template DNA concentrations.

2.6.1. TAQ POLYMERASE-BASED PCR

25µl reaction volumes were used for non radioactive PCRs where short length fragments (up to 1.5 kb) were to be amplified. The reaction contained template DNA of optimum concentration as previously assessed (see Section 2.6), a tenth volume of 10× Taq buffer (containing 50mM Tris-HCl (pH 8.0 at 25°C), 100mM sodium chloride, 0.1 mM EDTA, 1mM MDTT, 50% glycerol, and 1% Triton X-100), 0.2mM of dNTPs, 0.25µg of 5' and 3' PCR primers and 1.25 units of Taq polymerase. Amplification generally used 35 cycles of 95°C for 30 seconds, published or calculated primer-specific annealing temperature (see Section 2.6.3) for 45 seconds, and 72°C for 1-2 minutes according to product length. To decrease the possibility of primers misannealing, manual "hot starts" were used in many cases by

adding the enzyme in a buffered 5µl volume at the first annealing step without stopping the PCR programme. Hot starts were not continued at temperatures higher than the extension temperature of 72°C for any particular cycle to prevent evaporation.

10µl reaction volumes were used for radioactive PCRs in linkage analysis. PCR product sizes no larger than 280 base pairs were studied. A tenth volume of 10× Taq buffer (as above) was used with 0.5 units of Taq, 0.2mM of dNTPs, and 40ng of labelled 5' and unlabelled 3' PCR primers. Amplification used similar steps to those used in 25µl reactions but only 30 cycles were performed with the final extension step prolonged by 5 minutes before ramping to 0°C. Hot starts were not used as they were not required and would have increased the risk of radioactive contamination.

2.6.2. TTH/VENT POLYMERASE-BASED PCR

Unlike the tris-based Taq polymerase buffer where its acidity increases with increasing temperatures leading to breaks in the DNA being amplified, Tth/vent polymerase buffer, which is tricine based, remains neutral. This, coupled with the “proof reading” activity of vent, allows PCR amplification of much longer fragments to be achieved. PCR products of 2.7 kb were easily obtained in this work from cloned and genomic template DNA. Tth/vent based reactions were therefore used in most of the promoter analysis in order to amplify long sections. 25µl reactions were used with all reactions employing a manual hot start with 0.6 units of rTth DNA polymerase XL (Perkin Elmer) in a buffered 5µl volume. The same technique was

used as for Taq hot starts. 0.33 of final volume of $\times 3.3$ Tth XL buffer II (containing tricine, potassium acetate, glycerol, and DMSO) was used along with 0.1mM of dNTPs, 1 μ M of magnesium acetate, and 0.125 μ g of 5' and 3' primers. Similar amplification programmes to those for Taq based reactions were used with longer extension times used (up to three and a half minutes) reflecting the longer fragments to be amplified.

2.6.3. PRIMERS

Primers for PCRs used in the *endoglin* promoter analysis were designed using the promoter sequence (GeneBank accession number AF035753) (see Appendix C). Oligonucleotides were between 21 and 24 base pairs in length with no more than 60% of the bases being guanine or cytosine. Since these bases have 3 hydrogen bonds compared to the 2 hydrogen bonds of adenine and thymine, PCR is less efficient as their percentage increases. This is because the melting temperature, and thus the annealing temperature, of a primer is dependent on the number of hydrogen bonds that it contains; the more hydrogen bonds the higher the melting temperature. It was ensured that there were no complementary sequences within the primer and that primer ends were not complementary either to each other or to those of the other primer in the primer pair. The melting temperature for the primer was calculated as $4 \times (G+C) + 2 \times (T+A) ^\circ C$. Primers in a pair were designed to have similar melting temperatures. The annealing temperature used in a PCR programme was calculated as 5 $^\circ C$ less than the lower of the two primer melting temperatures. The calculated annealing temperature for a primer was not used if it was higher than the extension temperature. A lower one was used, but did not seem to lead to misannealing as

frequently as when shorter primers are used at inappropriately low annealing temperatures. Finally the designed primer was checked by the BLAST programme from The National Center for Biotechnology Information (NCBI), website <http://www.ncbi.nlm.nih.gov/>, to ensure that no similar sequences existed in the human genome to which it could misanneal.

Primers for PCR used in the linkage work had all been previously published^{231 232} (Appendices D and E).

2.7. DNA AMPLIFICATION: CLONING

2.7.1. INSERT PREPARATION AND LIGATION

PCR products were cloned. Products from Tth/vent polymerase based reactions were changed from having blunt ends to cohesive ends (see Section 2.5.2). The products were then ethanol precipitated and the concentration of product assessed by running a tenth on an agarose gel against a known marker. This allowed optimal insert: vector molar ratios (1:1 and 3:1) to be calculated for 50ng of a commercial vector, pGEM-T Vector (Promega). Vector and insert were then ligated (see Section 2.5.2). A control reaction containing no insert was also performed and used as a control in transformation.

2.7.2. TRANSFORMATION

Plasmid and insert DNA (2µl of 10µl ligation reaction) were added to 50µl of high efficiency competent cells (JM 109 High Efficiency Competent Cells (Promega)), handling the very fragile cells carefully, and placed on ice for 20 minutes. To allow DNA insertion into the cells, they were heat shocked at 42°C for 45-50 seconds and then iced for at least 2 minutes. 950µl of room temperature SOC medium was added and the cells allowed to grow at 37°C for one and a half hours. The transformation mixture was plated onto previously prepared LB agar plates at a variety of concentrations and cultured at 37°C overnight. The plates were spread with ampicillin, using 1.25mg/plate, to select for cells with plasmids which contained an ampicillin resistance gene. To identify colonies containing the inserts they were also spread with Xgal (0.75mg/plate) and 1M Isopropylthio-β-D-galactoside (IPTG) for blue-white selection. White colonies generally contained inserts. Inserts could also

be present in blue colonies, however, as in-frame inserts allow read-through and β -galactosidase activity. Identification of colonies containing the insert was also performed in some cases by colony PCR. 10 μ l Tth/vent polymerase based reactions (see Section 2.6.2) were performed using part of a colony (picked up by cocktail sticks) as template DNA and primers to amplify the insert in question. Colonies of interest were then grown overnight and the DNA extracted by alkaline lysis (see Section 2.2.2).

2.8. SEQUENCING

This was performed by thermal cycle sequencing for products amplified both by PCR and cloning. The Stratagene Cyclist Taq DNA sequencing kit was used. A 30µl premix was prepared from double distilled water, 4µl of 10× buffer, 16ng of an oligonucleotide primer end-labelled with Redivue³³P γ-ATP (see Section 2.5.3), 5 units of Taq polymerase, and template DNA. Template DNA included 2.5µl of non-purified PCR product and 3µl of ethanol precipitated digested cloned product. 6.7µl of premix was dispensed to each of four termination tubes containing 3µl of the respective termination mixes (ddGTP, ddATP, DDT, dicta). Reactions were prepared on ice and transferred to the thermal cycler (PTC 100 MJ Research) when the block temperature exceeded the annealing temperature. Sequencing programmes used were typically 35 cycles of 95°C for 20 seconds, the calculated annealing temperature of the specific primer (see Section 2.6.3) for 20 seconds, 72°C for 20 seconds. Samples were then stored at -20°C or immediately size-separated by polyacrylamide gel electrophoresis after denaturing (see Section 2.4.1.2). After detection by autoradiography the sequences were read by eye.

2.9. LINKAGE

2.9.1. ASCERTAINMENT OF PHENOTYPE AND GENOTYPE

Disease status was assigned using the current international consensus criteria²²⁹(see Section 2.1.3). Individuals who were designated status “unknown” were excluded from linkage analysis. Apparently unaffected children under the age of 12 were excluded as penetrance is less than 80% at this age³⁵. Linkage studies were performed by analysis of short tandem repeat polymorphic markers. Markers of high heterozygosity were chosen on the basis of their position relative to the genes of interest and experience in molecular papers (*Endoglin*^{35 136}; *ALK1*^{137 139}). Primers used to amplify the markers had been previously published^{231 232} (Appendices D and E). The markers were amplified from the extracted genomic DNA from each individual using 40ng of labelled Redivue ³³P γ-ATP 5’primer and 40ng of unlabelled 3’primer in a 10µl Taq polymerase based PCR reaction (see Section 2.6.1). The PCR products were denatured, size separated by denaturing polyacrylamide gel electrophoresis, detected initially by autoradiography, and, in some cases, further by phosphorimager. Allele patterns were assessed by eye by myself and a second observer (CLS), both of us being blinded to disease status.

2.9.2. LINKAGE ANALYSES

I performed two-point linkage analyses using the programme MLINK²¹⁵ via Human Genome Mapping Project (HGMP) computational resources. Based on previously published estimates³⁴ and previous studies of the families studied this time³⁵, penetrance was set at P=0.8 between ages 12 and 35 and P=0.95 for individuals aged

over 35 giving two liability classes. A disease gene frequency of 0.0001 was used, given that the prevalence of HHT in the population investigated is approximately 1 in 10,000 (see Section 1.1.2), and recombination rates in both sexes were assumed to be equal. Two point lod scores were calculated initially with equal allele frequencies and subsequently according to the allele frequencies in the population.

2.9.2.1. MULTIPOINT ANALYSES

The entered data for family S was subsequently compiled into two files (one including data from markers *D9S60*, *D9S315*, and *D9S61*, the other including data from markers *D12S85*, *D12S347*, *D12S368*, and *D12S325*) for multipoint analyses using GENEHUNTER²¹⁵, again via HGMP computational resources.

2.9.2.2. HAPLOTYPE ANALYSES

Haplotype analysis of chromosome 9 in Family P was performed by initially converting the text data files for *D9S60*, *D9S315*, *D9S61* and *HXB* to a format suitable for haplotype analyses using Mega2 via HGMP computational facilities²³³. Simwalk2 via the HGMP computational facility²³⁴ was then used to assess the number of possible haplotype solutions.

3. ENDOGLIN PROMOTER

Hypothesis : Mutations in the endoglin promoter region could account for disease in endoglin-linked families

Not all mutations have been found in *endoglin*-linked families studying all coding regions and splice sites of the gene³⁷. Given that haploinsufficiency is the prevailing model in *endoglin*-linked families^{37 148}, it seems likely that mutations in the promoter region of the gene could account for disease.

The 8 HHT families identified for this work were chosen as 4 had previously been shown to be linked to *endoglin* but no mutations had been found by Southern blotting, cycle sequencing of all known exons (starting 35 base pairs upstream of the transcription start site using endo1F) and intron/exon junctions, and cycle sequencing of entire cloned DNA³⁷ (starting 85 base pairs downstream of the transcription start site). No mutations had been found in the other 4 families using the same techniques but they were too small to allow significant linkage to an HHT gene to be made³⁷.

3.1. AMPLIFICATION OF SEQUENCE

A 3.2 kb sequence, encompassing the *endoglin* promoter and 605 base pairs spanning *endoglin* exon 1 and the start of the first intron, was kindly sent by a collaborator, Carmelo Bernabeu. At the time it had not as yet been published (it was published 10 months later¹³³). Primers were designed to amplify the entire 3.2 kb sequence using Tth/vent based reactions (see Figure 3.1) (see Section 2.6.2 and 2.6.3)(see Appendix C). Several different primer pairings were tried using DNA clones from a P1 library of human genomic DNA³⁷ as the target DNA. The DNA clone spanned the entire *endoglin* gene. Primer pairing GW1 (starting 10 base pairs downstream of the 5' end of the sequence and replicating the sense strand (sense primer)) and GW7 (2.7 kilobases downstream in exon 1 and replicating the antisense strand (antisense primer)) had reasonable success. A large amount of PCR product using this primer pair was obtained when the target DNA was cloned DNA. Restriction endonuclease digest of the PCR product using *Ava*I and *Pvu*II confirmed that the correct region had been amplified (see Figure 3.3). Previously prepared DNA from one affected individual from each of the 6 smaller families and from 2 affected individuals from both of the 2 larger families (both of which were linked to *endoglin*) was then used as target DNA. PCR product was obtained with difficulty with primer pairing GW1 and 7 (see Figure 3.2) and it was not reliable enough to be used exclusively. Smaller regions of the sequence were therefore amplified using different primer pairings. Good PCR product, 1.7 kb in length, was obtained with primer pair GW1 and GW3 (see Figure 3.4) using genomic target DNA. Amplification of the sequence starting immediately 3' to the region amplified by GW1 and 3 and extending to the start of

exon 1 was far more difficult. This was particularly frustrating as the main regulatory element was suspected (in fact correctly so¹³³) to be in this region of the sequence since it lay immediately upstream of the previously published exon 1 sequence. The difficulty with amplification and the importance of the region were probably not coincidental. Sequences for transcription factor binding sites and response elements in this region will occur in many other parts of the genome making specific amplification of this region alone problematic. Several different primer pairings were tried including nested primers on PCR products. Using sense primer GW11, starting 469 base pairs upstream of the transcription start site, with primer GW7, the region -469/+154 (with respect to the major transcription start site) was amplified. This was only achieved when 3 µl of PCR product obtained with primers GW11 and 7 was used as the target DNA for a further PCR using GW11 and 7 involving a total of 60 PCR cycles. This was not ideal because as the number of PCR cycles increases, the likelihood decreases of any polymorphism found being genuine rather than due to a replication error during PCR amplification.

3.2. SEQUENCING

An attempt to clone the PCR product spanning the entire 2.7 kilobase region between primers GW1 and 7 was made as this would allow cloned sequencing to be performed on a single product. However cloning this large fragment was not successful. Instead DNA from all individuals stated above which had been amplified with GW1 and 7, or GW1 and 3 and GW11 and 7, was cycle sequenced. Sequencing throughout the region of these PCR products was commenced (see Figure 3.1). Using a 5% longranger gel GW3 was used to sequence 273 base pairs at the 3' end of gel purified GW1 and 3 PCR product. No polymorphisms were found. GW4, starting 25 base pairs upstream of GW3 but acting as a sense primer unlike GW3, was used to sequence 390 base pairs from GW1 and 7 PCR product again using a longranger gel. This in fact encompassed the region -927/-666 later described as an Alu sequence¹³³. Again no polymorphisms were seen. GW8, designed to be a nested primer to GW7 as it started 23 base pairs upstream of GW7 and acted in the same direction, was used to sequence the 3' end of the GW1 and 7 product. GW8 however misannealed to another region of the product producing interesting results (see Section 3.3). Further sequencing was then focussed to the region immediately upstream of the major transcription start site. PCR product, from DNA from all individuals as described above, using primers GW11 and 7 was analysed. This was the region -469/+154¹³³. It was fully sequenced in both directions using GW11 and GW9 (204 base pairs downstream of GW11) as sense primers and GW7, GW15 (63 base pairs upstream of GW7), and GW12 (263 base pairs upstream of GW7) as antisense primers. No polymorphisms were found in any individual with any primer.

Different groups have included different proportions of the untranslated sequence 5' to the ATG start codon in their mutational screens. More extensive sequence was spanned in mutational studies by Shovlin et al³⁷ in which primer endo1F was derived from sequence 5' to the major transcriptional start sequence. This is 385 base pairs upstream of the ATG start codon as described by McAllister et al¹³⁶ and Yamaguchi et al¹⁴⁷ (who studied coding sequence only) (See Figure 3.5). Thus, although the region between the 5' untranslated region of exon 1 as described by Shovlin et al³⁷, based on the GenBank sequence X72012, and coding sequence as described by Rius et al¹³³, based on the GenBank sequence AF035753, was not screened in my work, it had previously been screened, in the families in my work, by Shovlin et al³⁷.

3.3. FORTUITOUS OUTCOME OF PRIMER GW8

MISANNEALING

As mentioned above, primer GW8 was designed to sequence the 3' end of the GW1 and 7 product as it started 23 base pairs upstream of GW7 and was an antisense primer. The sequence obtained with it however was completely different to the expected sequence. When the sequence obtained was checked by the BLAST programme from The National Center for Biotechnology Information (NCBI), website <http://www.ncbi.nlm.nih.gov/>, it was found to be identical to sequence just downstream of GW1. Knowing the site of the sequence that had been obtained, it was possible to deduce the region where GW8 must have misannealed. In the expected site, using comparison by eye, ten base pairs (TGTCAGTGCC) had identical sequence to that of the area for which GW8 had been designed, suggesting that this was likely to be the position of misannealing. The site of misannealing was the antisense strand starting 199 base pairs downstream of the 5' end of the GW1 and 7 fragment (see Figure 3.1). The region from the site of the misannealing to the 5' start of the GW1 and 7 PCR product had therefore been sequenced with this primer.

The results of the misannealing were interesting. One individual, individual 54 from family D, one of the families which was too small to allow linkage to *endoglin* to be made³⁷, showed dual sequence compared to the other individuals (see Figure 3.6). This appeared to be due to a 2 base pair deletion in one of the alleles. When sequencing the same region in the other direction using primer GW5, a nested primer to GW1 starting 9 base pairs downstream of GW1, dual sequence was again seen just in individual 54, again due to a 2 base pair deletion. Without separating the two

alleles for individual 54 it was difficult to pinpoint the exact site of the deletion although it was thought to be within 65 base pairs of the 5' end of the GW1 and 7 fragment. Therefore a small PCR product of 128 base pairs (126 base pairs if the deletion was present) was produced using primers GW1 and GW10 which encompassed the region of interest. This was successfully cloned (see Figure 3.7)(see Section 2.7) thus separating the two alleles. The cloned product was sequenced with primer GW10. This showed a 2 base pair CA deletion in 2 of the 10 clones, 46 and 47 base pairs respectively downstream of the 5' end of the GW1 and 7 fragment i.e. 2564 and 2563 base pairs upstream of the transcription start site (See Figure 3.8 and Figure 3.9).

3.4. FAMILIAL SEGREGATION OF POLYMORPHISM

To determine whether or not this CA deletion polymorphism cosegregated with disease in this pedigree, DNA from the other individuals in the family was then analysed (see Figure 3.10). A small PCR product encompassing the site of the deletion was obtained for individuals 53, 55, and 56, as well as further product for individual 54. Unfortunately the only way to obtain product was from a nested reaction using DNA amplified by primers GW1 and 10, as target DNA for a further 50 µl Tth/vent reaction using a nesting sense primer GW5, starting 9 base pairs downstream of GW1, and primer GW10. This meant that 60 PCR cycles in total were used. The presence or absence of the deletion in each individual was then assessed by analysing the PCR product by restriction endonuclease digests and cycle sequencing.

3.4.1. RESTRICTION ENDONUCLEASE DIGESTS TO ANALYSE FAMILY D

Restriction endonuclease digests of the GW5/10 PCR product for all family D individuals was performed (see Figure 3.11) with XbaI to confirm that the correct PCR product was being analysed and with BsrI to detect the deletion.

3.4.1.1. XBAI RESULTS

The GW5/10 product should be digested into fragments of 76 and 44 base pairs irrespective of whether or not the deletion is present. At first sight (see Figure 3.11) the expected digestion appeared to have occurred. However, on further inspection, the larger fragment was seen to be different in size in different individuals making the results impossible to interpret.

3.4.1.2. BSRI RESULTS

The deletion should abolish the BsrI restriction site so that a complete undigested 118 base pair product would be seen. If no deletion were present, digestion would occur giving two fragments of 83 and 37 base pairs in length. As can be seen in the gel in Figure 3.11, it appeared that digestion occurred in individual 53 but not in the other individuals suggesting that individual 53 was the only member of the family without the polymorphism. This again was inconclusive however because in individual 54, known to be heterozygous for the polymorphism and therefore having one wild type allele instead of two, no digested wild type allele was observed. This implied that the digestion might have failed. Alternatively, it is possible that the mutant allele had been preferentially amplified by the PCR. The GW5 and 10 PCR products for all family D individuals were therefore then cycle sequenced instead.

3.4.2. CYCLE SEQUENCING TO ANALYSE FAMILY D

The GW5 and 10 PCR products for all family D individuals were sequenced using GW16, a primer nested to GW10 starting 15 base pairs upstream of GW10 and acting on the antisense strand. This demonstrated that individuals 54, 55, and 56 had the deleted and normal alleles whereas individual 53 only had the normal alleles. This meant that the CA deletion was passed from the affected mother (individual 54) to the affected daughter (individual 56). However the deletion was also seen in the unaffected spouse of the daughter (individual 55) and so this indicated that it must be a polymorphism not a disease-causing mutation (see Figure 3.12).

3.5. TRANSCRIPTION FACTOR BINDING SITES DISRUPTED BY THE CA DELETION

To assess whether or not this polymorphism could have any effect on function, potential transcription factor binding sites which could be encoded by this region, and so be disrupted by the deletion, were searched for using computer programmes designed to test for transcription factor binding sites. Using MatInspector V 2.2, high similarity was found between this region and sequence encoding the retroviral oncogene viral myb, and the transcription factor gut-enriched Kruppel-like factor (GKLF). However other programmes, designed using more recent consensus sequence, did not detect the same similarities. Close inspection of the sequences from primary references indicated that there was insufficient homology for function to be possible.

3.6. ALLELE FREQUENCY OF THE POLYMORPHISM

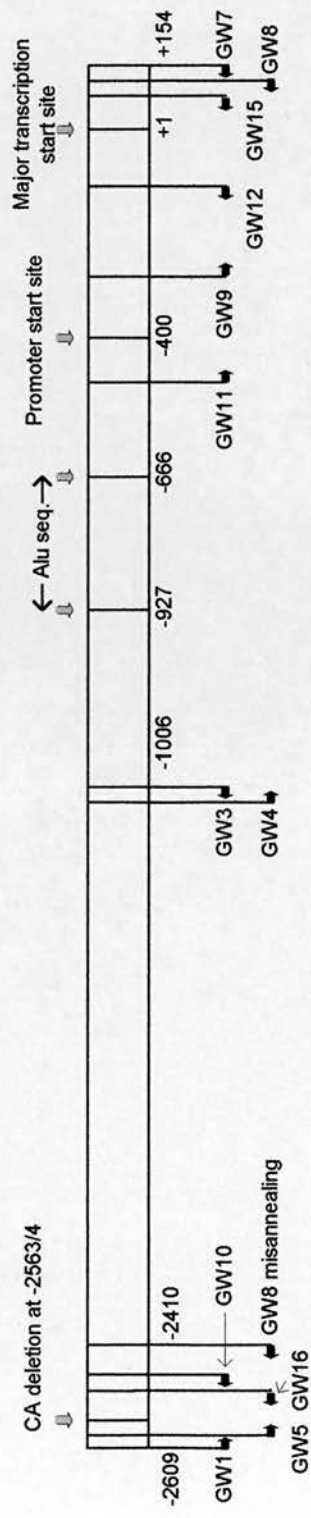
To determine the allele frequency in the UK population of this polymorphism, 100 normal chromosomes were analysed. DNA from 50 unaffected spouses was amplified in the same way as previously with PCR product from a GW1 and 10 primed reaction being used as target DNA for a GW5 and 10 primed reaction. The product was sequenced with GW16. The polymorphism was seen in 8 of the 50 individuals, occurring in both alleles in one of them. This gave an allele frequency of 9% for this polymorphism (see Figure 3.13).

3.7. SUMMARY

In the families analysed no mutations in the *endoglin* promoter region were found that could account for disease in *endoglin*-linked families. The promoter region was fully sequenced in both directions in affected individuals from the 8 families identified with no mutation or polymorphism seen. This cannot fully discount the hypothesis that mutations in the *endoglin* promoter region could account for disease in *endoglin*-linked families as only a small number of families were analysed.

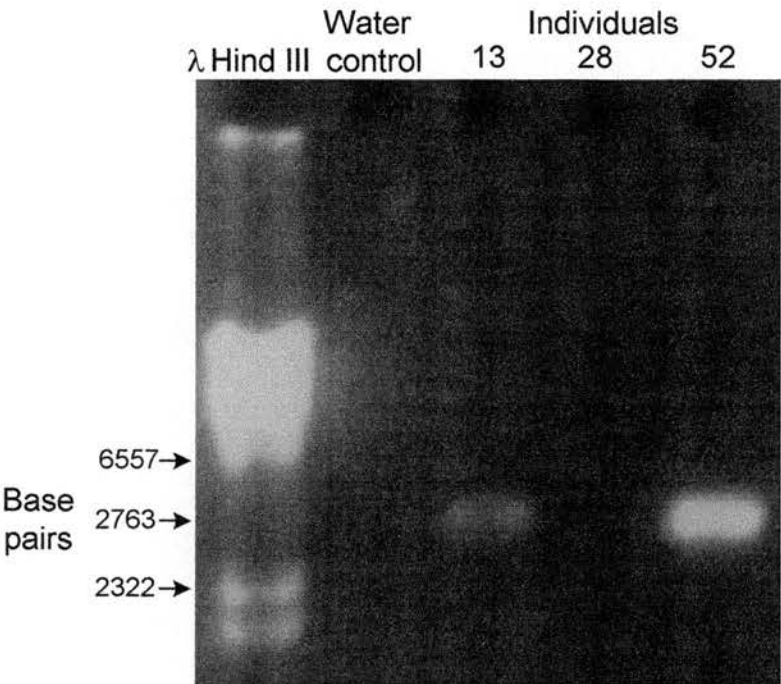
A two base pair CA deletion was found 2563 base pairs upstream of the transcription start site in an affected individual from one of the families. It was found to cosegregate with disease but was also seen in the unaffected spouse of the affected daughter indicating that it was a polymorphism rather than a mutation. It has been shown to have an allele frequency of 9% in the UK population.

**FIGURE 3.1 REGION AROUND THE *ENDOGLIN* PROMOTER AND
THE POSITION OF PRIMERS USED**



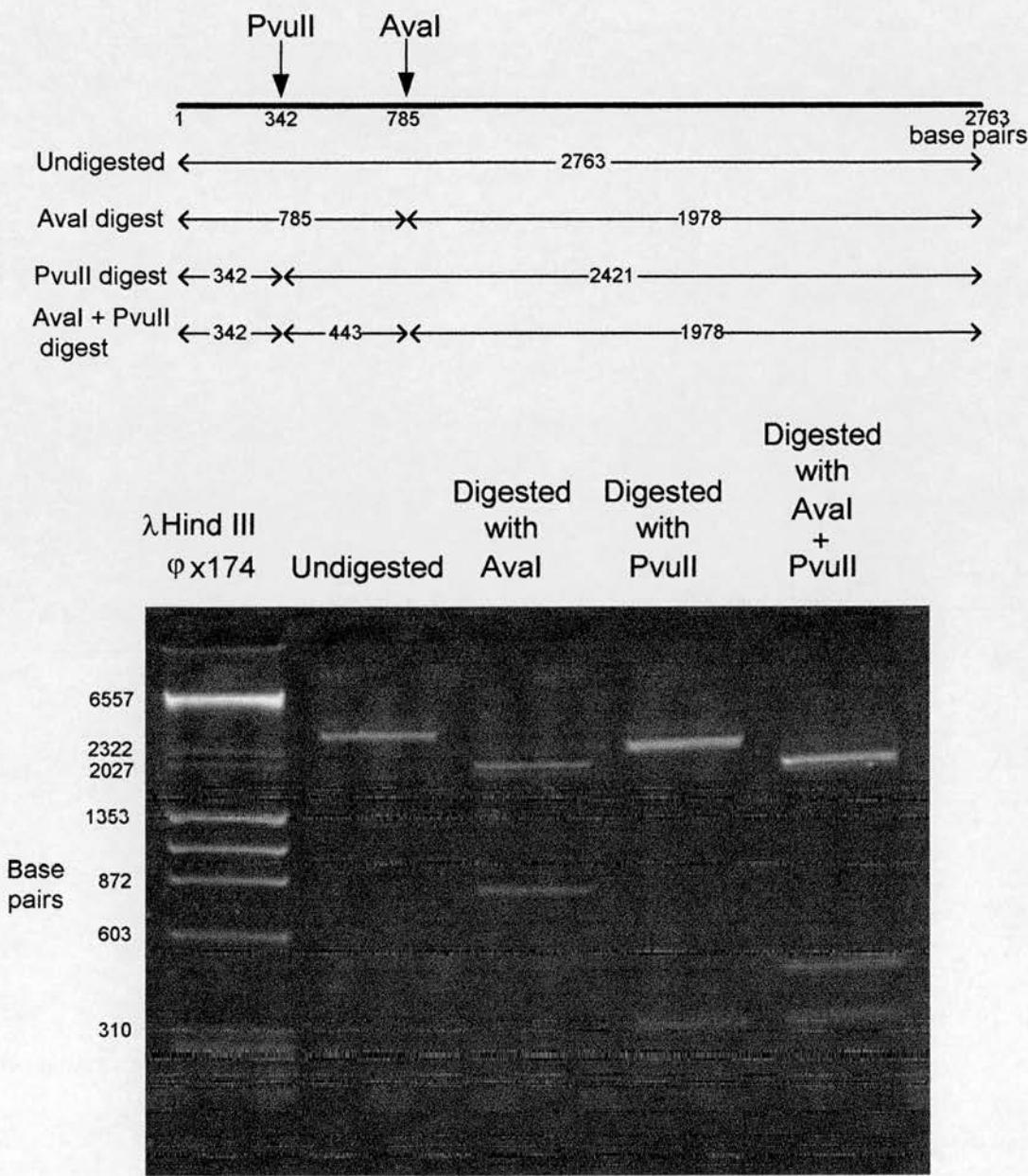
Primers used, and direction they act in, are indicated by solid black arrows. The number below the long box represents the base pair number with respect to the major transcription start site at base pair +1.

**FIGURE 3.2 PCR PRODUCTS USING PRIMERS GW1 AND GW7 AND
GENOMIC TARGET DNA**



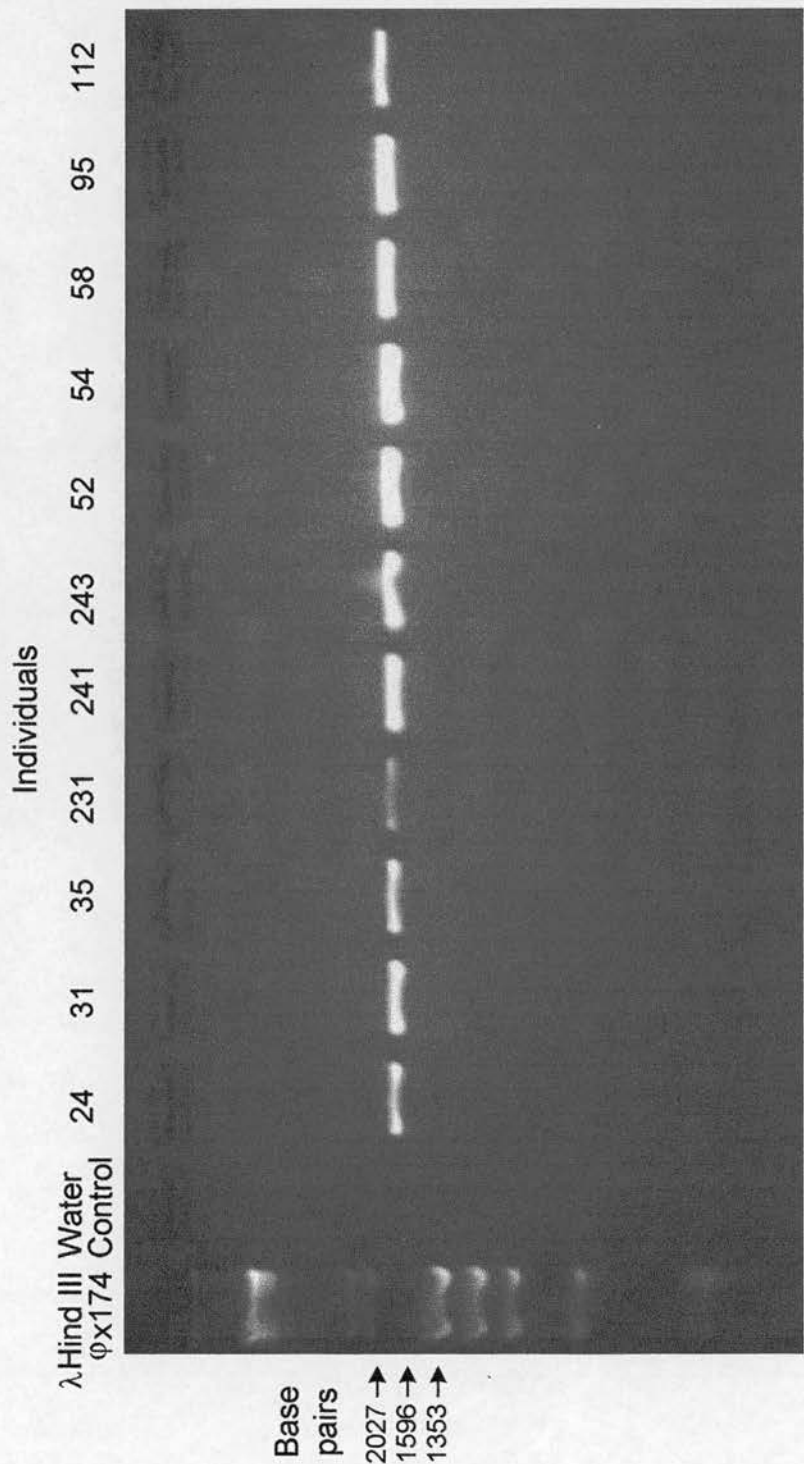
PCR product 2763 base pairs in length was obtained with primers GW1 and 7 and genomic target DNA from affected individuals.

**FIGURE 3.3 RESTRICTION ENDONUCLEASE DIGEST WITH AVAI
AND PVUII OF PCR PRODUCT USING PRIMERS GW1
AND GW7 ON CLONED GENOMIC *ENDOGLIN***



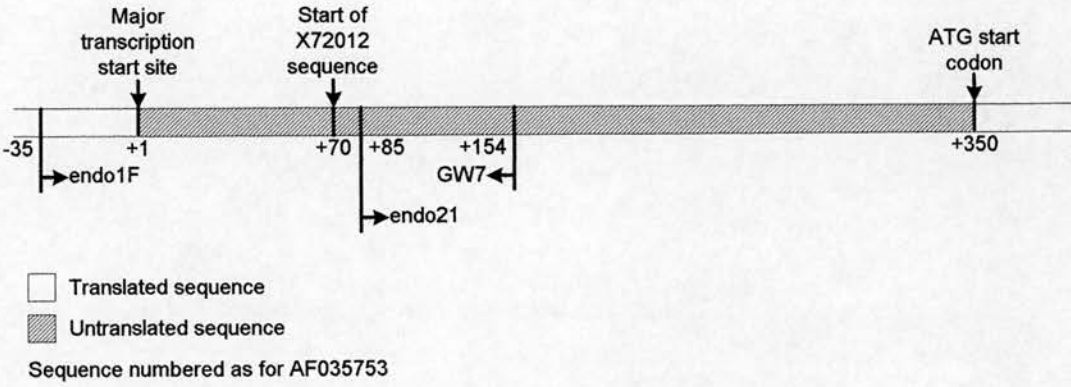
The expected fragment lengths were obtained when GW1 and 7 PCR product was digested with Aval and PVUII. This confirmed that GW1 and 7 had amplified the region of DNA for which they had been designed.

FIGURE 3.4 **PCR PRODUCTS USING PRIMERS GW1 AND GW3 AND GENOMIC *ENDOGLIN***



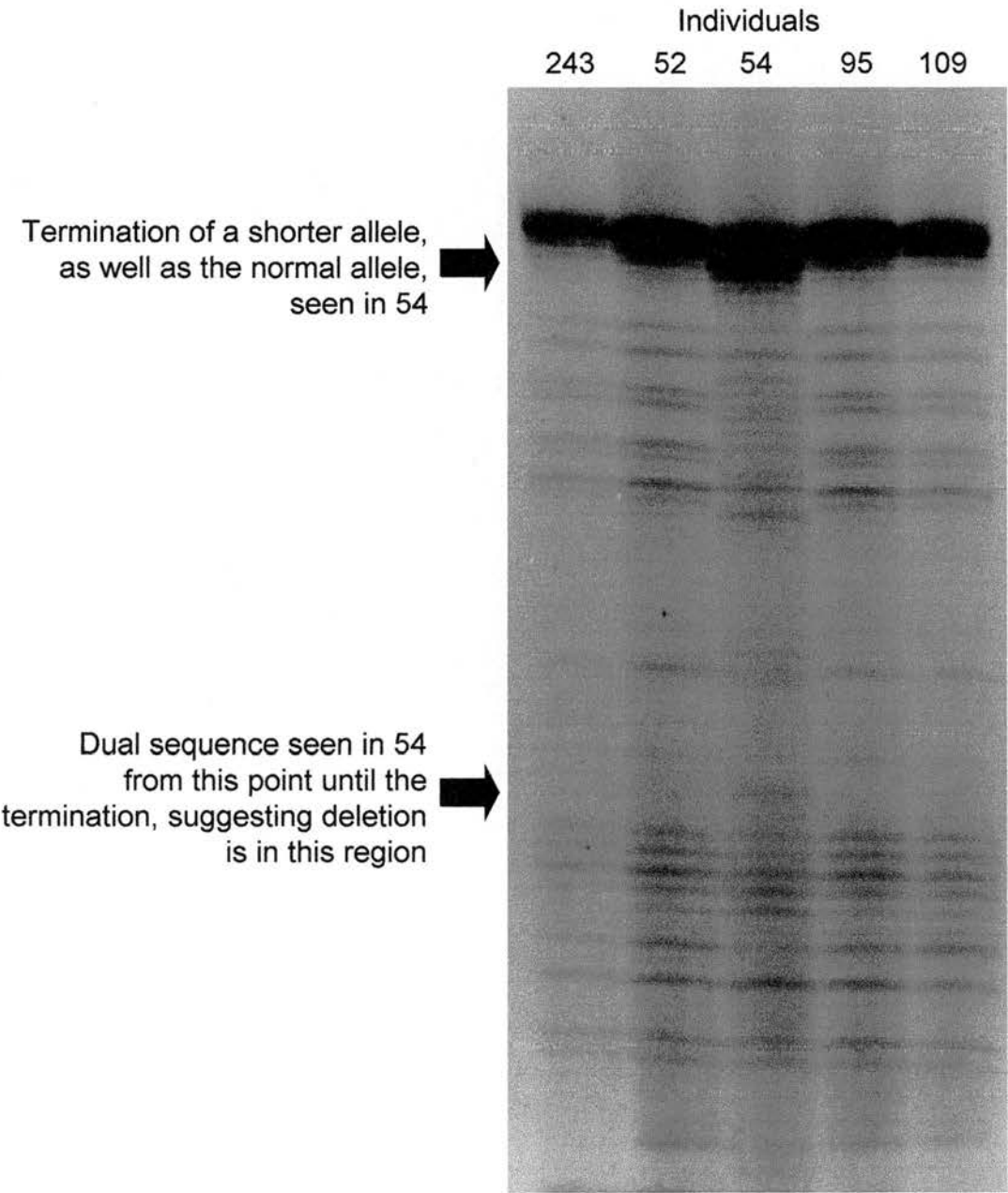
PCR product 1596 base pairs in length was obtained with primers GW1 and 3 and genomic target DNA from affected individuals.

**FIGURE 3.5 COMPARISON OF GENBANK SEQUENCES AF035753
AND X72012**



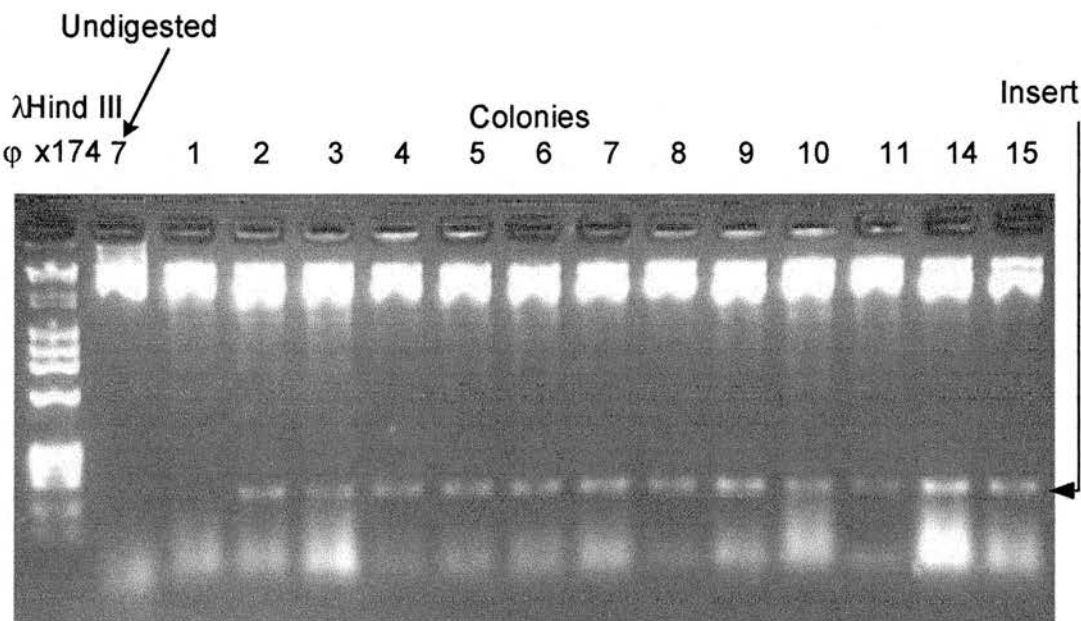
Primer endo1F was derived from sequence 5' to the major transcriptional start sequence and used in mutational studies by Shovlin et al³⁷. Primer GW7 was the most 3' primer used in my work. Thus, although the region between the 5' untranslated region of exon 1 as described by Shovlin et al³⁷, based on the GenBank sequence X72012, and coding sequence as described by Rius et al¹³³, based on the GenBank sequence AF035753, was not screened in my work, it had previously been screened, in the families in my work, by Shovlin et al³⁷.

FIGURE 3.6 THE G TRACK OF SEQUENCE USING PRIMER GW8



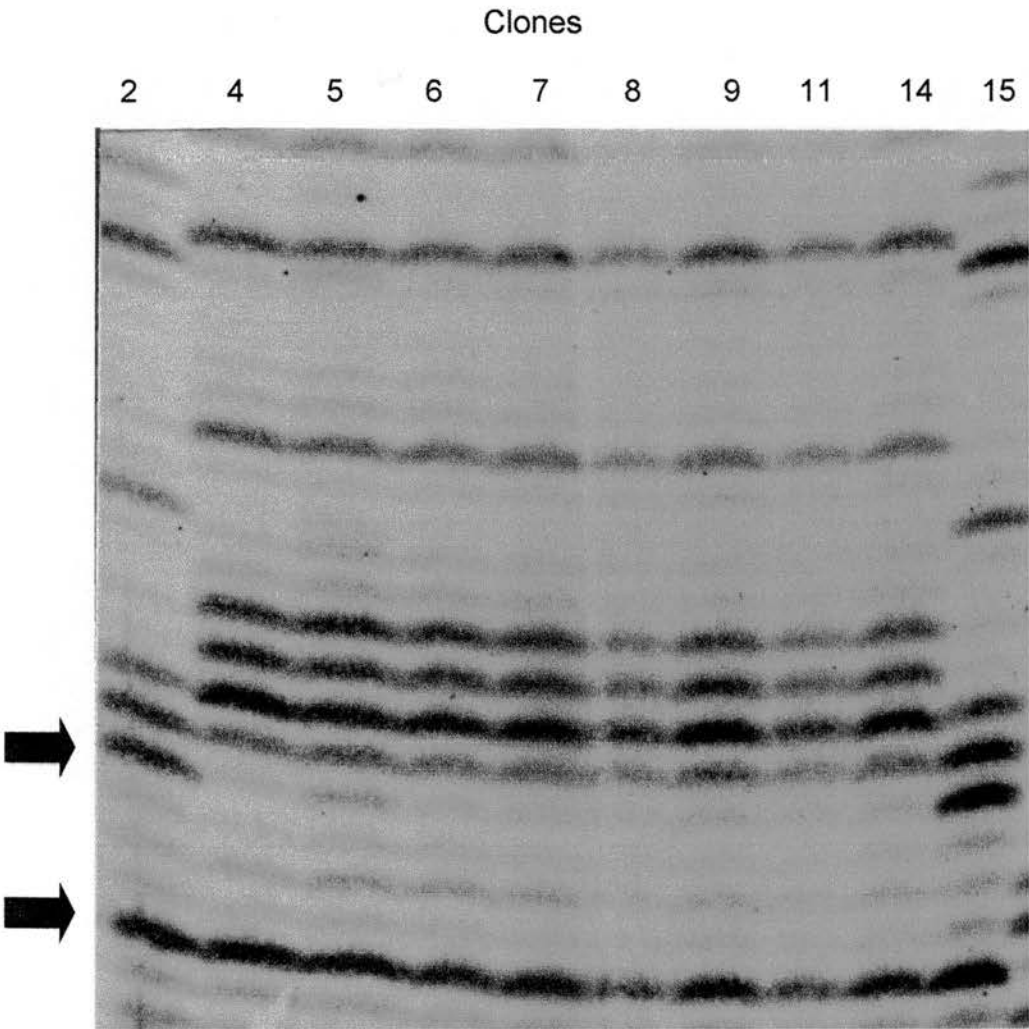
Dual sequence was seen in affected individual 54 from family D when cycle sequencing with primer GW8 on GW1 and 7 PCR product.

**FIGURE 3.7 COLONIES DIGESTED WITH ECORI TO OBTAIN
CLONED GW1/10 INSERT**



A small PCR product of 128 base pairs was produced using primers GW1 and 10 on DNA from individual 54. This amplified the region where the deletion had been detected previously. The product was cloned to separate the 2 alleles and the cloned inserts were obtained by digestion with EcoRI.

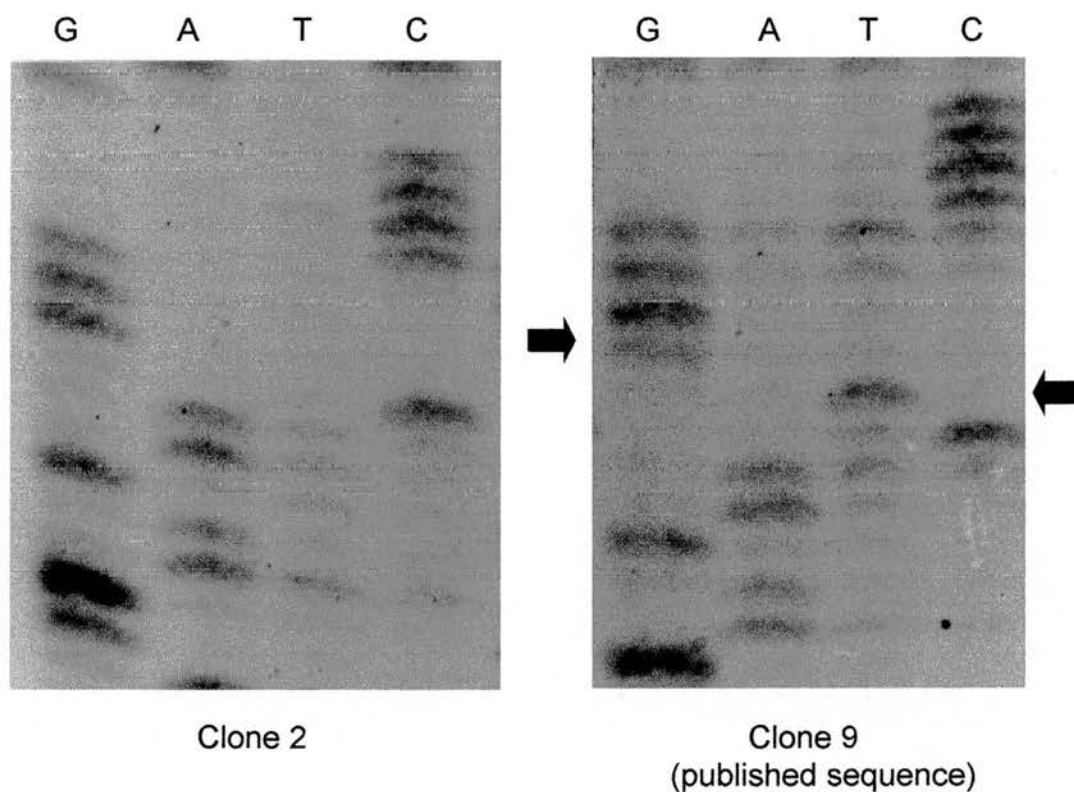
FIGURE 3.8 THE G TRACK OF 10 SEQUENCED CLONES



2 base pair deletion, one of which
is a g, occurring in clones 2 and
15 between the arrows

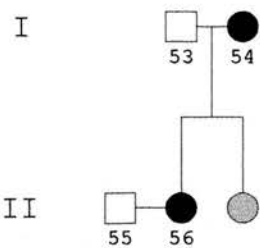
GW1 and 10 PCR product from affected individual 54 was cloned and the clones sequenced with primer GW10. Two of the clones, numbers 2 and 15, showed a 2 base pair deletion.

**FIGURE 3.9 SEQUENCING GELS SHOWING G AND T BASES
ABSENT IN CLONE 2**



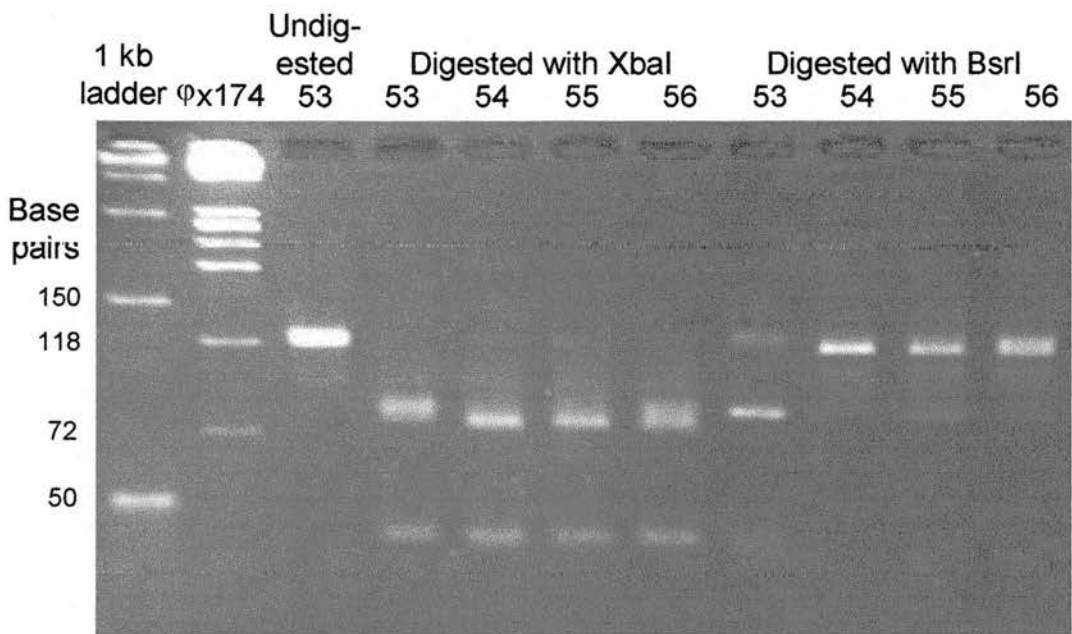
GW1 and 10 PCR product from affected individual 54 was cloned and the clones sequenced with primer GW10. Two of the clones, numbers 2 and 15, showed a 2 base pair deletion. Looking at all four nucleotide tracks reveals that the deletion in clone 2 is due to an absent T and G base (marked with the solid black arrows in clone 9). As the antisense strand was sequenced this is a CA deletion in the published sense strand.

FIGURE 3.10 FAMILY TREE OF FAMILY D



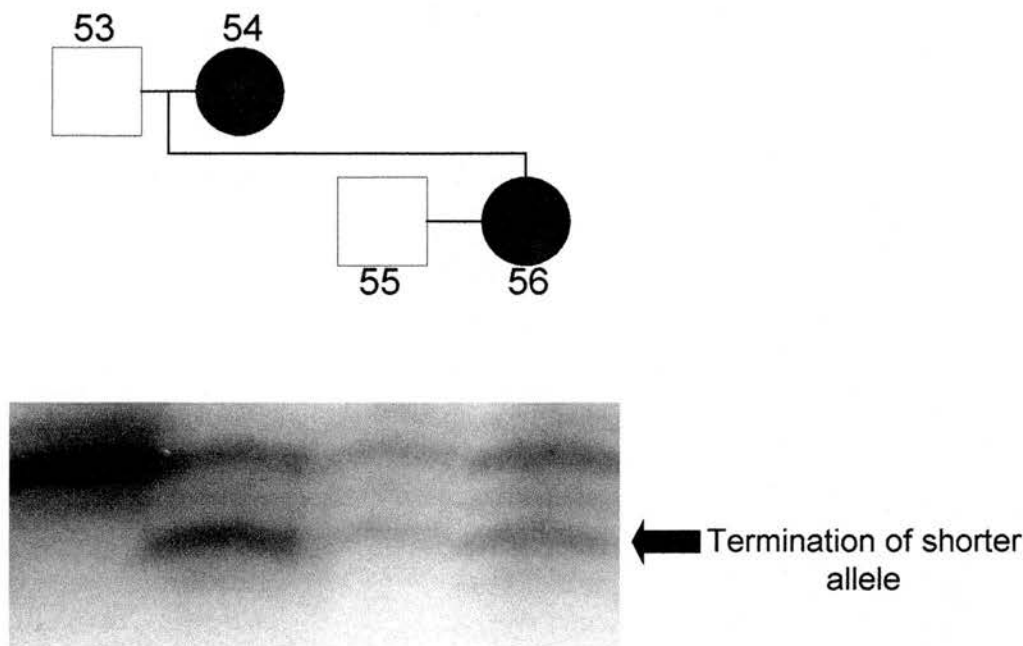
The numbered family members are the ones which were analysed. Affected individuals are indicated by black symbols, unaffected by white symbols, and the individual with unknown disease status is in grey. Squares represent males and circles represent females.

**FIGURE 3.11 RESTRICTION ENDONUCLEASE DIGEST WITH XBAI
AND BSRI OF PCR PRODUCT USING PRIMERS GW5
AND 10 ON DNA FROM MEMBERS OF FAMILY D**



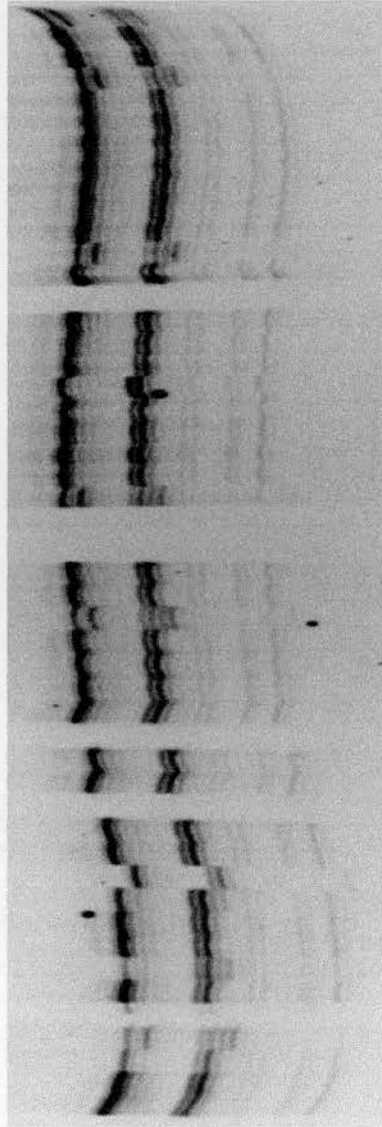
The region where the deletion had been found in individual 54 was amplified with GW1 and 10 and then GW5 and 10 in other members of family D. The product was digested with XbaI to confirm that the correct PCR product was being analysed. It was then digested with BsrI. If the deletion was present it should abolish the BsrI restriction site preventing digestion (see Section 3.4.1 for further details).

FIGURE 3.12 FAMILY D SEQUENCING GEL



The region where the deletion had been found in individual 54 was amplified with GW1 and 10 and then GW5 and 10 in other members of family D. The product was sequenced with GW16, a primer nested to GW10. This picture shows the termination of the C track of sequence for each member of family D. Dual sequence was seen in individuals 54, 55, and 56 with termination of a shorter, as well as a normal length, allele. This indicated that the deletion had been passed from the affected mother to the affected daughter but was also present in the unaffected husband of the daughter. The deletion was therefore a polymorphism not a disease-causing mutation.

**FIGURE 3.13 TERMINATIONS IN G TRACK SEQUENCING WITH
GW16 ON GW5/10 PCR PRODUCT TO ASSESS ALLELE
FREQUENCY OF POLYMORPHISM**



The allele frequency in the UK population of the CA deletion polymorphism was calculated by amplifying DNA from 50 unaffected spouses with PCR product from a GW1 and 10 primed reaction being used as target DNA for a GW5 and 10 primed reaction. The product was sequenced with GW16. The polymorphism was seen in 8 of the 50 individuals (termination of a shorter allele as well as the normal allele is demonstrated in this gel), occurring in both alleles in one of them. This gave an allele frequency of 9% for this polymorphism.

4. LINKAGE

Hypothesis: HHT families with pulmonary involvement may be unlinked to the PAVM gene, endoglin, and may be unlinked to both known HHT genes, endoglin and ALK1

This hypothesis was suggested by two pieces of evidence. First, the work of an Italian group suggests that a third gene for HHT could exist¹⁴⁰, although no pulmonary involvement was seen in the pedigree described by this group. Secondly PAVMs have been shown not to occur exclusively in *endoglin*-linked families^{35 37}, although *ALK1* status was unknown.

As stated previously families S and P, initially identified in 1992 when one individual from each presented to the Hammersmith Hospital for management of PAVMs, were assessed further by myself and CLS six years later. They were considered as being suitable for testing the above hypothesis as they had pulmonary involvement and initial analyses in 1992 suggested that they were unlinked to *endoglin*, but there were insufficient numbers to allow *ALK1* linkage analyses³⁵. Clinical evaluation, venepuncture, DNA extraction, and linkage analyses were performed for each individual as described previously.

4.1. FAMILY S

4.1.1. RESULTS OF CLINICAL EVALUATION

The pedigree for family S is shown in Figure 4.1 with the results of the clinical investigations in Table 4.1. Family members who were not assessed are not shown. Table 4.2 shows that overall 11 new diagnoses were made, 3 in original members who had now reached phenotypable age, and 8 from the 10 additional members assessed. Four generations were assessed with 12 affected members identified from all generations. Individual IV.3 was given a diagnosis of obligate affected as she had affected offspring and also an affected parent, but her only manifestation was epistaxis. Individuals V.1, V.5, and V.11 had epistaxis only but no affected children and so were classified of unknown status and excluded from further analysis. Individual IV.3 had a geographical and fish tongue.

4.1.1.1. FAMILY S MEMBERS WITH A DIAGNOSIS OF PAVMS

As mentioned in Section 2.1.2.1, individuals were suspected of having PAVMs based on haemoptysis, central cyanosis, or clubbing, SaO₂ erect $\leq 96\%$ (as recommended by¹¹⁴), or supine-erect fall greater than 3.3% (thus exceeding 2SDs beyond the mean of 0.3% (SD1.5) in the non-PAVM group in¹¹⁴). Individual V.3 fulfilled these criteria by having an erect SaO₂ of $\leq 96\%$ and individual V.4 fulfilled the criteria by having a supine-erect fall in SaO₂ of 4%. Individual IV.2 also fulfilled the criteria by having an erect SaO₂ of $\leq 96\%$. He declined further investigation for PAVMs however.

4.1.1.2. FURTHER DETAILS OF FAMILY S MEMBERS WITH A DIAGNOSIS OF PAVMS

Individual V.3 was the proband for this family, and had presented aged 23 years with a long-standing history of painless haemoptysis and a one year history of episodic dizziness and blurred vision. He was not dyspnoeic. Examination revealed bilateral early clubbing and a bruit was audible over the right lower lobe. He was polycythaemic (Hb 18.9, haematocrit 0.55), and hypoxaemic (pH 7.44, pCO₂ 4.5kPa, pO₂ 8.2kPa) with additional orthodeoxia (supine SaO₂ 92%, erect 87%). Spirometry was normal (FEV₁ 5.25 litres, FVC 5.95 litres) but the gas transfer was markedly reduced with TLCO 6.96 (52% predicted), KCO 1.87 (51% predicted). Chest X-rays and a thoracic CT scan confirmed the presence of several PAVMs. The right-to-left (R-L) shunt measured by ^{99m}Techetium-labelled albumin macroaggregates was 24% of the cardiac output. Following selective right pulmonary angiography, two large PAVMs were embolised with 21 steel coils resulting in immediate improvement in oxygen saturations (supine SaO₂ 97%, erect 96%). His initial family screening dates from this period. The improvement in SaO₂ was sustained two years post embolisation. The R-L shunt remained at 4.2%, probably accounted for by a residual right basal PAVM with feeding arteries too small to embolise. When we assessed him 6 years post embolisation his dizzy spells had recurred, oxygen saturations fell to 96% supine, 95% erect. His haematocrit was mildly raised at 0.564 (normal range 0.40-0.54) but his haemoglobin was normal at 165g/l. He is due to be re-evaluated with further pulmonary angiography.

Individual V.4 had been first evaluated as part of a family screen (by CLS) at the age of 21 years when she had no epistaxes or telangiectasia and normal oxygen saturations (SaO₂ 100% in supine and erect postures). When reviewed by CLS and myself for this study, she was aged 27 years, and was 10 months post partum. She described the development of monthly epistaxes and frequent dizzy spells after her pregnancy. On examination, she was not clubbed or cyanosed, but she demonstrated significant desaturation in the erect posture (SaO₂ 99% supine, 94% erect). Her subsequent review five months later by respiratory physicians, following our recommendation to her general practitioner, indicated that her hypoxaemia had persisted (SaO₂ 97% supine, 93% erect, 92% post 90 watts of exercise). Her haematocrit was raised at 0.515 (normal range 0.37-0.47) but her haemoglobin was normal at 148g/l. Spirometry was normal (FEV₁ 4.00 litres, 117% predicted; FVC 4.3 litres, 109% predicted), but the gas transfer was marginally impaired (KCO, corrected for Hb, 1.59, 88% predicted). Contrast echocardiography indicated the presence of a R-L shunt, though a thoracic CT scan (axial scanning at 5mm intervals) did not reveal any macroscopic PAVMs. A diagnosis of microscopic PAVMs was made. On review six months later, her oximetry showed spontaneous improvement at rest (SaO₂ 98% in supine and erect postures) but post-exercise oximetry and gas transfer remained marginally impaired (SaO₂ 96% following 120 watts of exercise; corrected KCO 1.55, 85% predicted), and radionucleotide scanning quantitated the R-L shunt at 7.2% of the cardiac output (normal < 3.5%). She remains on prophylactic antibiotic therapy and under respiratory review, currently in her second pregnancy.

4.1.2. LINKAGE ANALYSES

4.1.2.1. EXCLUSION OF *ENDOGLIN*

The position of *endoglin* relative to the polymorphic markers used is shown in Figure 4.2. The *HHT1* locus was originally mapped to a 2cM interval on chromosome 9 flanked by *D9S60* and *D9S61*¹³⁵. The adjacent *D9S61*-*D9S63* interval was suggested by other investigators¹³⁵ but later revised by the same group to between *D9S60* and *D9S61*¹³⁶. The endoglin gene has been mapped physically as equidistant between *D9S315* and *SPTAN* which both lie within the interval flanked by *D9S60* and *D9S61*. The relative positions of *endoglin* and *ALK1* have not been resolved. Markers *D9S60*, *D9S315*, and *D9S61* were initially used.

Two point lod scores obtained with markers *D9S60*, *D9S315*, and *D9S61* were sufficient to exclude linkage to *endoglin* in family S (see Figure 4.3). Using the definition of a score of -2 or less indicating no linkage at a significance level of 0.05, these data exclude linkage to a region of about 15 cM on both sides of the *endoglin* locus. This is also demonstrated by multipoint analysis (see Figure 4.3).

4.1.2.2. EXCLUSION OF *ALK1*

Analysis of linkage to *ALK1* was performed. Figure 4.4 shows the location of *ALK1* relative to its flanking polymorphic markers. Two original assignments of the *HHT2* locus were given, one between *D12S339* and *D12S359*¹³⁷, and the other in a neighbouring interval between *D12S345* and *D12S339*¹³⁸ (marked with a broken line in Figure 4.4). The latter was revised on subsequent analyses and the *ALK1* gene is now known to lie in the 4cM interval between *D12S347* and *D12S359*

(marked with a solid black line with grey extension in Figure 4.4), most likely in the 1cM interval between *D12S347* and *D12S368*¹³⁹(marked with a solid black line in Figure 4.4). Nevertheless I considered it prudent to examine both disease intervals using markers *D12S331*, *D12S85*, *D12S347*, and *D12S368* initially.

In family S marker *D12S368* did not exclude linkage to the region flanked by it and *D12S325* (see Figure 4.5), although the region between it and *D12S331* was excluded. Marker *D12S325* was therefore analysed. Lod scores obtained for it were sufficiently negative to exclude linkage to the *D12S368-D12S325* region. Two point lod scores therefore excluded *ALK1* linkage in family S and this exclusion was further confirmed by multipoint analysis for these markers in family S (see Figure 4.5).

4.1.3. ALLELE FREQUENCIES

Lod scores were initially analysed using equal allele frequencies. Some of the markers were reanalysed using altered allele frequencies reflecting the frequencies seen in the pedigree studied. This did not significantly alter any of the lod scores obtained (see Table 4.3).

4.1.4. HAPLOTYPE ANALYSES

Limited haplotype analyses in families S were performed.

In family S, on both chromosome 9 and chromosome 12, of the four affected individuals in generation IV (three of whom were parents or grandparents of affected individuals so must have been correctly phenotyped), two displayed one haplotype which could be traced to their affected parent and two the other. Analysis of the

haplotypes transmitted to and by the individuals with proven PAVMs (V.3, V.4), who clearly also have the correct disease status assignment, was particularly instructive (see Figure 4.6). It can be seen that on both chromosomes 9 and 12, both V.3 and V.4 have inherited the haplotype (B) from their affected mother that she inherited from her unaffected father (with a recombination event between *D12S85* and *D12S347* in V.4 on chromosome 12). Moreover, on both chromosomes, V.3 has transmitted to his affected daughter (VI.1) the haplotype that he inherited from his unaffected father (C). This confirms that HHT in this family cannot result from mutations in either *endoglin* or *ALK1*.

4.1.5. SUMMARY

Family S is an HHT-PAVM family unlinked to the two known HHT genes, *endoglin* and *ALK1*.

4.2. FAMILY P

4.2.1. RESULTS OF CLINICAL EVALUATION

The pedigree of family P is shown in Figure 4.7 along with the results of the clinical investigations in Table 4.4. Family members who were unavailable for assessment are not shown. As shown in Table 4.2, 2 new diagnoses were made in original members due to these members reaching phenotypable age, and 20 additional members were assessed enabling 17 further new diagnoses. In total, 19 new diagnoses were made. 24 members spanning 4 generations were affected by HHT as defined by the consensus criteria²²⁹. Individuals III.1, IV.9, and IV.10 were classified as obligate affected as they all had affected offspring although personally fulfilled only two diagnostic criteria. Individuals V.4 and V.10 had epistaxis only and no affected children. They were categorised as being of unknown status and further analysis was not performed. Individual IV.13 had no epistaxis. She had possible telangiectasia but they were difficult to distinguish from spider telangiectasia and she had a history of heavy alcohol usage. Her son, individual V.15 had several episodes of epistaxis each year but this was possibly traumatic in nature. He had only a possible telangiectasia and no affected offspring. They were both therefore also classified as being of unknown status and not analysed further. It would be interesting to obtain a CT head scan for individual V.15 given his history of epilepsy, as this could possibly be due to HHT related CAVMs. Strictly speaking, however, even if a diagnosis of CAVMs was confirmed, this would still not be enough to allow disease status to be assigned to him or his mother²²⁹. Individual V.3 had a large telangiectatic lesion on her thumb during the third trimester of her

pregnancy which improved markedly after the delivery of her daughter. It is possible that her pregnancy had also worsened or precipitated PAVMs. Her oxygen saturations were worse postpartum probably since the pregnancy had masked PAVM associated hypoxaemia. She had had normal oxygen saturations 6 years previously. In contrast individual IV.5 demonstrated the often deleterious effect that a low oestrogen state has on HHT since her epistaxis had become more severe, and she had started to have GI bleeds (although the cause is unknown), since the menopause. Individual IV.5 had a fish tongue and individual V.7 had a geographical tongue. These features have not been associated with HHT previously and could, of course, be incidental, but it is interesting to note that one of the affected individuals from family S (individual IV.3) had a geographical and fish tongue.

4.2.1.1. FAMILY P MEMBERS WITH POSSIBLE PAVMS

The same criteria for PAVM screening (see Section 2.1.2.1) as when screening family S members were used to screen family P members. Individual V.3 had an erect SaO₂ of $\leq 96\%$ post partum and individual V.5 also had an erect SaO₂ of $\leq 96\%$ and was clubbed. Individual III.1 was suspected of having PAVMs on clinical examination alone (clubbing) and individuals V.7, VI.1, and VI.3 were suspected on clinical history alone (haemoptysis). Further investigation for PAVMs was therefore recommended for all of these: some declined follow up and the remainder are under review by local respiratory physicians.

4.2.2. LINKAGE ANALYSES

4.2.2.1. INCONCLUSIVE *ENDOGLIN* LINKAGE DATA

The position of *endoglin* relative to the polymorphic markers used is shown in Figure 4.2. The *HHT1* locus was originally mapped to a 2cM interval on chromosome 9 flanked by *D9S60* and *D9S61*³⁵. The adjacent *D9S61-D9S63* interval was suggested by other investigators¹³⁵ but later revised by the same group to between *D9S60* and *D9S61*¹³⁶. The *endoglin* gene has been mapped physically as equidistant between *D9S315* and *SPTAN* which both lie within the interval flanked by *D9S60* and *D9S61*. The relative positions of *endoglin* and *AK1* have not been resolved. Markers *D9S60*, *D9S315*, and *D9S61* were initially used.

Lod scores obtained for markers *D9S60*, *D9S315*, and *D9S61* did not suggest linkage to the *endoglin* locus between *D9S315* and *D9S61* (if the disease gene in this family were linked to *endoglin*, markers *D9S60*, *D9S61* and *D9S315* would be expected to show strongly positive lod scores at low recombination values at $\theta < 0.02$ ³⁵).

However, in contrast to Family S, positive lod scores were obtained with some of these markers (see Table 4.5). The significantly positive lod scores obtained with marker *D9S61* could not be interpreted as excluding linkage to *endoglin*. In particular the score of 1.27 at $\theta = 0.2$ with marker *D9S61*, indicated that the regions 20-30 cM from the region spanned by the three markers should be analysed further. Further markers were therefore studied to clarify this (see Table 4.6). Significantly positive lod scores were again obtained with marker *HXB*, but there was no consistent pattern to confirm or refute linkage to the region.

4.2.2.1.1. HAPLOTYPE ANALYSES

Given the inconclusive data obtained for family P when assessing *endoglin* linkage using two point analyses, I attempted to obtain additional information by performing haplotype analyses of chromosome 9 markers. Data was converted using Mega 2²³³ and analysed using Simwalk2²³⁴ both via HGMP computational facilities. Limited analyses using *HXB*, *D9S60*, *D9S315*, and *D9S61* indicated that there were three potential haplotype solutions preventing the derivation of additional data by this method. Figure 4.8 illustrates allele inheritance for the purposes of discussion of allele frequencies in the pedigree.

4.2.2.1.2. MODELS CONSTRUCTED FOR INCONCLUSIVE *ENDOGLIN* LINKAGE DATA

Three possible models were constructed for this inconclusive data: firstly that family P was linked to *endoglin* and that the method did not detect this linkage, secondly that the disease gene lay approximately 20-30cM away, and thirdly that the family was unlinked to both *endoglin* and neighbouring markers and spuriously high lod scores had been obtained with some of the markers.

4.2.2.1.2.1. MODEL 1: ANALYSES HAVE FAILED TO CONFIRM LINKAGE IN AN *ENDOGLIN*-LINKED FAMILY

If the disease gene in this family were linked to *endoglin*, markers *D9S60*, *D9S315*, and *D9S61* would be expected to show strongly positive lod scores at low recombination values of $\theta < 0.02$ ³⁵. Negative lod scores at low values of θ hypothetically could have arisen by one of a number of potential limitations of linkage analyses as discussed in the introduction (see Section 1.5.5). Specifically:

1) Misphenotyping- The negative lod scores within 1-2 cM of the *endoglin* gene could have resulted from misphenotyping of genetically unaffected individuals as affecteds, as in this setting the MLINK program assigns a probability of 1.0 that the assigned phenotype is correct. This is in contrast to an unaffected diagnosis which is made with <100% confidence (defined in our age-related penetrance model as being at probabilities of 0.8 at ages 12-35 and 0.95 at ages >35). Probands IV.3, IV.5, IV.7 and V.8 are concordant across *D9S60-61*, but IV.1 could only be concordant at *D9S60* (see Figure 4.8). As *endoglin* lies between *D9S315* and *D9S61*, that would imply IV.1 and descendants (V.2, VI.1, and VI.2), III.1 and descendent (IV.12) and V.7 would all have had to be misphenotyped. This possibility was considered carefully, particularly as none of these individuals had documented visceral involvement (although of note VI.1 had haemoptysis when running, III.1 was clubbed, and V.7 had haemoptysis daily for 4 months) (see Table 4.4). It must be noted that III.1 did not have epistaxis and was designated obligate affected due to having telangiectasia and an affected offspring (IV.12) according to the current international consensus criteria²¹⁷. All of the other probands had epistaxis and they all, including III.1, had telangiectasia. While spontaneous nosebleeds can occur in non-HHT patients, a non-HHT cause of classical HHT telangiectasia cosegregating with disease in this family to result in these misdiagnoses would be highly unlikely, but remains hypothetically possible. Therefore misphenotyping to generate artificially negative lod scores could only have occurred if we were to postulate the unlikely scenario of a second condition (genetic or environmental) contributing to these clinical features in these probands.

2) Incorrect assignment of parentage- Although it is hypothetically possible that non-parentage could have contributed, this would require the affected father of IV.1 to be different to the affected father of IV.3, IV.5, and IV.7, and neither of these fathers to be the brother of III.1. IV.7 would also not be the father of V.7 who would have a different affected father. This is highly unlikely. In any case there was no evidence of non-parentage in this family as in all markers studied expected Mendelian segregation was seen (see Section 1.5.5.2).

Conclusion: This model does not appear likely. Furthermore it should be noted that *endoglin* mutational analysis has been conducted on 2 probands in the pedigree (IV.7 and V.8) and indicated no evidence for mutations in the genomic DNA (Southern blotting, cycle sequencing of all known exons and introns or junctions) or cloned DNA³⁷ supporting the conclusion that *endoglin* is unlikely to be the disease gene in this family.

4.2.2.1.2.2. MODEL 2: IN FAMILY P THERE IS A NOVEL HHT DISEASE GENE SITED APPROXIMATELY 20-30CM AWAY FROM THE *ENDOGLIN* LOCUS

The significantly positive lod scores obtained with marker *HXB* were not obtained with the other markers in the vicinity, markers *D9S261*, *D9S289*, and *D9S195*, which demonstrated lod scores of -2.30 at $\theta = 0.05$, -7.37 at $\theta = 0.00$, and -3.33 at $\theta = 0.01$ respectively (see Table 4.6). Normally, such findings would be taken to exclude linkage but the model in question demands that we should consider whether or not the analyses of *D9S261*, *D9S289*, and *D9S195* could be spurious. Possible causes for this could include:

- 1) Incorrect siting of markers on the genetic map so that distances are inappropriately low between these markers. This would have been a significant concern in the early days of linkage mapping, but is less likely given the extensive work by Genethon²³² and other investigators that went into siting these markers. It is particularly unlikely since *HXB* lies in the middle of the markers, between *D9S289* and *D9S195*. Since recombination mapping using Genethon pedigrees was used to site the markers, errors in order are significantly less likely than errors in distance, as this would require frequent double recombinations in small genetic distances between two adjacent markers.
- 2) Sample mix-up. A key feature of my laboratory training was to take judicious care never to mix up samples at any stage of DNA preparation or analysis. Nevertheless the possibility will be discussed to assess whether this was feasible based on the results. If samples had been mixed up when working with markers *D9S261*, *D9S289*, and *D9S195*, but not when working with marker *HXB*, this could produce the results obtained in this work. However this would mean that for marker *D9S261* the key “mixed up” samples would be for individuals III.1, IV.11, IV.12, and V.7 whereas for markers *D9S289* and *D9S195* samples for different groups of individuals would need to be mixed up (III.1, IV.1, IV.12, V.2, VI.1, and VI.2 for *D9S289*; III.1, IV.10, IV.12, V.7, and V.12 for *D9S195*) (see Figure 4.8). Furthermore it should be noted that it is often an individual plus the individual’s affected descendants whose samples would need to be mixed up rather than random individuals’ samples e.g. the samples for IV.1 plus all her descendants would need to be mixed up when marker *D9S289* was used. Sample

mix up is rendered even less likely as non-Mendelian patterns of inheritance of marker alleles would result and this was not seen.

Conclusion: From the data in this thesis alone, linkage to this region of chromosome 9 cannot be excluded. While *HXB* in itself is not a good candidate, there are other candidates in the region including the type 1 TGF β receptor, *ALK-5*. Further linkage studies are underway in the Shovlin laboratory and the results are awaited.

4.2.2.1.2.3. MODEL 3: THERE IS NO HHT DISEASE GENE IN THE REGION ON CHROMOSOME 9. THE *D9S61* AND *HXB* LOD SCORES ARE SPURIOUSLY HIGH AS A RESULT OF COMMON ALLELES PRESENT IN THE POPULATION APPARENTLY SEGREGATING WITH DISEASE

In this setting we should consider whether the lod scores for markers *HXB* and *D9S61* are artificially high. For *HXB* this possibility could arise if the allele shared at *HXB* by IV.1, IV.3, IV.5, and IV.7, (designated allele 1), is common in the general population and the affected father of IV.1, IV.3, IV.5 and IV.7 was homozygous for allele 1. Although the heterozygosity index for *HXB* is 0.79, it could be lower in this particular population. Since two of the ten spouses' alleles (from IV.2, IV.4, IV.6, IV.8, and V.1) are allele 1, this very limited survey indicates this might be a relatively common allele in the local population, and the possibility that the affected father of IV.1, IV.3, IV.5 and IV.7 was homozygous for allele 1 must be considered. The fact that his sister was homozygous for allele 5 does not exclude this possibility. Similar considerations can be applied to discussions of the positive lod scores with *D9S61*.

The possession of particular alleles by non-affected spouses was specifically analysed to investigate the possibility that artificially high lod scores might be generated by their bringing the allele designated as “disease” allele into the pedigree. For all the markers, except *GSN*, *ABL*, and *D9S66*, at least one of the five unaffected spouses carried the allele shared by affected individuals IV.3, IV.5, and IV.7.

Conclusion: Model 3 is plausible.

4.2.2.1.2.4. CONCLUSIONS FROM THE MODELS

It can be seen from these three models that the data remains inconclusive although the most likely model appears to be the third one- that the family is unlinked to both *endoglin* and neighbouring markers and spuriously high lod scores were obtained with some of the markers.

4.2.2.2. EXCLUSION OF *ALK1*

Analysis of linkage to *ALK1* was performed. Figure 4.4 shows the location of *ALK1* relative to its flanking polymorphic markers. Two original assignments of the *HHT2* locus were given, one between *DI2S339* and *DI2S359*¹³⁷, and the other in a neighbouring interval between *DI2S345* and *DI2S339*¹³⁸ (marked with a broken line in Figure 4.4). The latter was revised on subsequent analyses and the *ALK1* gene is now known to lie in the 4cM interval between *DI2S347* and *DI2S359* (marked with a solid black line with grey extension in Figure 4.4), most likely in the 1cM interval between *DI2S347* and *DI2S368*¹³⁹ (marked with a solid black line in Figure 4.4). Nevertheless I considered it prudent to examine both disease intervals using markers *DI2S331*, *DI2S85*, *DI2S347*, and *DI2S368* initially.

Two point lod scores obtained for family P (see Table 4.7) significantly excluded linkage to *ALK1*. Although marker *D12S325* was not analysed in this pedigree, results obtained with marker *D12S368* significantly excluded linkage to a region extending to about 18 cM from it and so easily excluded the region flanked by *D12S325*.

4.2.3. ALLELE FREQUENCIES

Lod scores were initially analysed using equal allele frequencies. Some of the markers were reanalysed using altered allele frequencies reflecting the frequencies seen in the pedigree studied (see Table 4.8). This did not significantly alter any of the lod scores obtained (see Table 4.9).

4.2.4. SUMMARY

Family P is an HHT family with probable pulmonary involvement which is unlinked to *ALK1* but the data regarding *endoglin* linkage is inconclusive.

4.3. FAMILIES V AND T- ASSESSMENT OF *ALK1* LINKAGE IN KNOWN NON-*ENDOGLIN* HHT-PAVM FAMILIES

Two additional HHT families with pulmonary involvement, families V and T, had previously been reported as being unlinked to *endoglin*³⁷ (No positive lod scores obtained for either family; family V: $Z = -2.58$, $\theta = 0.00$, *D9S315* ; family T: $Z = -2.06$, $\theta = 0.01$, *D9S61*). No mutations had been found by Southern blotting of the entire *endoglin* gene, cycle sequencing of all known exons and intron/exon junctions, and cycle sequencing of the entire cloned DNA³⁷. DNA obtained when the families were clinically assessed in 1992 was used for analysis. The families were not assessed further clinically by myself. The family trees for families V and T are shown in Figure 4.9. Individual III.1 from family V had a PAVM (25% R-L shunt at presentation with associated clubbing). Individual III.4 from family T had a symptomatic PAVM (R-L shunt 15.5% at presentation). I hypothesised that in these families HHT could result from either a mutation in the *ALK1* gene (not previously associated with HHT leading to the development of PAVMs) or that they could represent further families unlinked to the known HHT genes *endoglin* and *ALK1*. Six years ago they had also been assessed for *ALK1* linkage with no significant results obtained. However since then the disease interval has been refined allowing markers much closer to *ALK1* to be used and so increasing the likelihood of obtaining significant results. Markers *D12S347* and *D12S368* were used initially (see Table 4.10).

4.3.1. FAMILY V

Two point lod scores obtained with *DI2S368* exceed +1.2 over the interval spanning the *ALK1* locus, and a possible extension if earlier, now superseded work, was current¹³⁷⁻¹³⁹. This indicates significant linkage to the known *ALK1* locus in family V. In the course of this project, due to time constraints, I was unable to pursue this to identify the causative *ALK1* gene mutation.

4.3.2. FAMILY T

In family T the negative two point lod scores obtained with *DI2S368* suggested that this family could be a non-*endoglin*, non-*ALK1* HHT-PAVM family, but the scores with this marker, and the adjacent marker *DI2S339*, were insufficiently informative. In the course of these studies there was insufficient time to pursue these analyses further.

4.3.3. SUMMARY

Family V is an HHT-PAVM family linked to *ALK1*. Family T is an HHT-PAVM family unlinked to *endoglin* with inconclusive *ALK1* linkage data.

4.4. CONCLUSIONS

1. These results demonstrate that not all HHT families with pulmonary involvement are linked to the known HHT-PAVM gene, *endoglin*, and furthermore that one family, family S, is unlinked to the two known HHT genes, *endoglin* and *ALK1*.
2. The most important finding is that family S, an HHT-PAVM family, is unlinked to the two known HHT genes, *endoglin* and *ALK1*. Since *endoglin* and *ALK1* are both involved in the regulation of response to ligands of the TGF β superfamily, it seems likely that the gene mutated in family S may encode another component of the TGF β signalling complex or a downstream effector. Its identification should significantly enhance our understanding of signalling by TGF β family members. The finding also has an immediate clinical implication since it is clear from this work that all HHT families, not just *endoglin*-linked families, require PAVM screening.
3. The linkage of an HHT-PAVM family, family V, to *ALK1* indicates that PAVM screening should be performed in *ALK1*-linked families. This reinforces the implication regarding PAVM screening derived from the family S data.
4. The inconclusive nature of the analyses in family P highlights some of the limitations of linkage analyses. The modelling of theoretical possibilities was a useful exercise to critically analyse the methodology and physical basis of linkage studies.

TABLE 4.1 CLINICAL FEATURES OF FAMILY S

No.	Age	Epistaxis		Telangiectasia	PAVM Screen			Diagnosis
		Onset	Current severity		Clinical	SaO2(S)	SaO2(E)	
III.1	76	-	-	-	ANPS	97	97	Unaffected
III.2	69	childhood	occasional	buccal	ANPS	98	97	Affected
III.3	72	-	-	-	ANPS	-	-	Spouse
IV.1	45	-	-	-	ANPS	-	-	Spouse
IV.2	49	childhood	none	lips, facial		96	96	Affected
IV.3	48	childhood	none since cauterised aged 16	-	ANPS	98	97	Obligate affected
IV.4	50	-	-	-	ANPS	-	-	Spouse
IV.5	48	childhood	2 monthly	facial	ANPS	100	100	Affected
IV.6	48	-	-	-	ANPS	-	-	Spouse
IV.7	42	childhood	frequent, several cauterisations	lips, buccal	ANPS	98	98	Affected
IV.8	40	-	-	-	ANPS	-	-	Spouse
V.1	20	childhood	occasional	?buccal, ?arms	ANPS	97	97	Unknown
V.2	36	-	-	-	ANPS	-	-	Spouse
V.3	28	preschool	almost daily	buccal, tongue, fingers	previously embolised for PAVMs	96	95	Affected
V.4	27	aged 26	monthly	lips, hands	"dizzy spells"	98	94	Affected
V.5	22	childhood	none since childhood	-	ANPS	99	96	Unknown
V.6	19	childhood	occasional	buccal	ANPS	99	98	Affected
V.7	17	childhood	weekly	buccal	ANPS	99	98	Affected
V.8	9	aged 4	several each year	tongue, hand	ANPS	98	98	Affected
V.9	18	childhood	occasional	buccal	ANPS	98	98	Affected
V.10	17	-	-	-	ANPS	98	98	Unaffected
V.11	10	-	twice precipitated by trauma	-	ANPS	98	98	Unknown
VI.1	9	preschool	several each year	facial, fingers	ANPS	98	97	Affected

*=prepartum **=postpartum ***= no blood sample obtained

Clinical: positive features of respiratory history and examination

ANPS= asymptomatic, no physical signs S= supine E= erect

Pedigree numbers correspond to Figure 4.1. See text for additional clinical information about V.3 and V.4. Method for PAVM screening was adapted from Thompson *et al*¹¹⁴ in which SaO2 in the erect posture of $\leq 96\%$ gave a sensitivity of 73% and specificity of 35%, for detection of PAVMs as defined by selective pulmonary angiography, and in which the supine-erect SaO2 was significantly greater in the PAVM rather than non-PAVM group: 1.7% (SD 2.7) compared to 0.3% (SD1.5) $p=0.01$. Therefore SaO2 erect $\leq 96\%$ or supine-erect fall greater than 3.3% (thus exceeding 2SDs beyond the mean of 0.3% (SD1.5) in the non-PAVM group in ¹¹⁴) was used to suggest possible PAVMs requiring further investigation.

TABLE 4.2 DIAGNOSES OF HHT

Family	S	P
Total no. seen	23	36
No. of new members seen	10	20
Affected	11	21
Obligate affected	1	3
Unaffected	2	3
Unknown	3	4
Unaffected spouse	6	5
New Diagnoses in Original Members	3	2
New Diagnoses in New Members	8	17

Summary of number of family members of families S and P clinically assessed and diagnoses made.

**TABLE 4.3 ANALYSIS OF LOD SCORES WITH EQUAL AND
ADJUSTED ALLELE FREQUENCIES IN FAMILY S**

Family	Locus	Allele freq.	Recombination fraction (θ)							
			0.00	0.01	0.05	0.10	0.20	0.30	0.40	0.50
S	<i>D9S315</i>	equal	-16.07	-7.67	-4.93	-3.21	-1.57	-0.72	-0.25	0.00
S	<i>D9S315</i>	adjusted	-15.95	-7.60	-4.91	-3.19	-1.56	-0.72	-0.24	0.00
S	<i>D12S347</i>	equal	-9.07	-5.14	-4.60	-3.89	-2.16	-1.14	-0.47	0.00
S	<i>D12S347</i>	adjusted	-9.06	-5.14	-4.59	-3.80	-2.10	-1.13	-0.47	0.00

No significant difference was found between lod scores obtained using equal and adjusted allele frequencies.

TABLE 4.4 CLINICAL FEATURES OF FAMILY P

No.	Age	Epistaxis		Telangiectasia	PAVMs			Other possible HHT manifestations	Diagnosis
		Onset	Current severity		Clinical	SaO2 supine	SaO2 erect		
III.1	66	-	-	nasal, lip, buccal	clubbed	97	97	TIA's	Obligate affected
IV.1	61	? Childhood	none	lips, buccal	ANPS	97	97	-	Affected
IV.2	62	-	-	-	ANPS	-	-	-	Spouse
IV.3	58	childhood	occasional	facial, buccal, hands	ANPS	97	97	-	Affected
IV.4	62	-	-	-	ANPS	-	-	-	Spouse
IV.5	54	childhood	daily	lips, buccal, hands	ANPS	97	98	upper GI bleeding, fish tongue	Affected
IV.6	51	-	-	-	ANPS	-	-	-	Spouse
IV.7	50	teenage	daily despite laser treatment	facial	ANPS	98	98	-	Affected
IV.8	53	-	-	-	ANPS	-	-	-	Spouse
IV.9	55	-	-	lip	ANPS	97	98	-	Obligate affected
IV.10	42	-	-	lip	ANPS	97	98	-	Obligate affected
IV.11	37	childhood	several each year	facial, lips, buccal, tongue	ANPS	97	97	-	Affected
IV.12	46	childhood	weekly	tongue	ANPS	98	97	bright blood per rectum	Affected
IV.13	45	-	-	round lips ?ethanol related	ANPS	96	96	-	Unknown
V.1	38	-	-	-	ANPS	-	-	-	Spouse
V.2	38	childhood	none since aged 10	buccal	ANPS	96	98	-	Affected
V.3	35	childhood	several each week	facial, buccal, hands		98*, 94/5**	97*, 93/2**	Thumb lesion 3rd trimester, resolved postpartum	Affected
V.4	34	childhood	several per year	-	ANPS	96	97	-	Unknown
V.5	32	childhood	twice weekly	facial	clubbed for 6 years	96	95	-	Affected
V.6	29	-	-	-	ANPS	98	98	-	Unaffected
V.7	23	aged 21	weekly	facial, lip, hands	haemoptysis daily for 4 months	99	98	geographical tongue	Affected
V.8	21	teenage	monthly	facial, fingers	ANPS	98	98	-	Affected
V.9	30	teenage	several each year	lips	ANPS	97	97	-	Affected
V.10	27	aged 22	twice monthly	-	haemoptysis once years ago	98	98	-	Unknown

Continued on next page.

No.	Age	Epistaxis		Telangiectasia	PAVMs			Other possible HHT manifestations	Diagnosis
		Onset	Current severity		Clinical	SaO2 supine	SaO2 erect		
V.11	26	-	-	-	ANPS	97	97	-	Unaffected
V.12	16	aged 15	monthly	lip	ANPS	99	99	-	Affected
V.13	13	-	-	-	ANPS	97	98	-	Unaffected
V.14	8	aged 3	several each year	buccal, hands, arms	ANPS	98	97	-	Affected ***
V.15	23	teenage	2-3 each year but precipitated by trauma	?top lip	ANPS	96	96	epilepsy	Unknown
VI.1	14	childhood	monthly	nasal, buccal	haemoptysis when running	97/6	97/5	-	Affected
VI.2	12	childhood	several each year	lip, hand	ANPS	97	97	migraines	Affected
VI.3	17	childhood	daily	hands	haemoptysis every 2 weeks	98	98	epilepsy but normal CT head scan	Affected
VI.4	11	childhood	several each year	lip, hands	ANPS	97	97	-	Affected
VI.5	9	childhood	monthly	buccal, arms	ANPS	98	97	-	Affected ***
VI.6	12	preschool	several each year	hands	ANPS	97	96	-	Affected ***
VI.7	9	childhood	several each year	hands	ANPS	97	97	-	Affected

*=prepartum **=postpartum ***= no blood sample obtained

Clinical: positive features of respiratory history and examination

ANPS= asymptomatic, no physical signs S= supine E= erect

Pedigree numbers correspond to Figure 4.7. Method for PAVM screening was adapted from Thompson *et al*¹¹⁴ in which SaO2 in the erect posture of $\leq 96\%$ gave a sensitivity of 73% and specificity of 35%, for detection of PAVMs as defined by selective pulmonary angiography, and in which the supine-erect SaO2 was significantly greater in the PAVM rather than non-PAVM group: 1.7% (SD 2.7) compared to 0.3% (SD1.5) $p=0.01$. Therefore SaO2 erect $\leq 96\%$ or supine-erect fall greater than 3.3% (thus exceeding 2SDs beyond the mean of 0.3% (SD1.5) in the non-PAVM group in ¹¹⁴) was used to suggest possible PAVMs requiring further investigation.

TABLE 4.5 **TWO POINT LOD SCORES FOR FAMILY P WITH**
MARKERS *D9S60*, *D9S315*, AND *D9S61* FLANKING THE
***ENDOGLIN* LOCUS**

	Locus		
Recombination Fraction (θ)	<i>D9S60</i>	<i>D9S315</i>	<i>D9S61</i>
0.00	-17.14	-11.86	-6.27
0.01	-4.88	-3.79	-2.28
0.02	-3.40	-2.68	-1.71
0.03	-2.54	-2.01	-1.15
0.04	-1.95	-1.53	-0.70
0.05	-1.49	-1.17	-0.34
0.10	-0.21	-0.08	0.67
0.20	0.64	0.72	1.27
0.30	0.72	0.81	1.17
0.40	0.45	0.54	0.71
0.50	0.00	0.00	0.00

Lod scores were calculated at different recombination fractions.

**TABLE 4.6 TWO POINT LOD SCORES FOR FAMILY P WITH
ADDITIONAL CHROMOSOME 9 MARKERS**

Locus	Recombination fraction (θ)							
	0.00	0.01	0.05	0.10	0.20	0.30	0.40	0.50
<i>D9S261</i>	-14.50	-5.76	-2.30	-1.82	-0.78	-0.32	-0.09	0.00
<i>D9S289</i>	-7.37	-0.77	0.47	0.83	0.90	0.67	0.35	0.00
<i>HXB</i>	1.48	1.45	1.33	1.17	0.81	0.44	0.11	0.00
<i>D9S195</i>	-17.17	-3.33	-0.32	0.69	1.17	0.94	0.44	0.00
<i>GSN</i>	-14.23	-6.58	-3.08	-1.61	-0.42	0.00	0.09	0.00
<i>D9S60</i>	-17.14	-4.88	-1.49	-0.21	0.64	0.72	0.45	0.00
<i>D9S315</i>	-11.86	-3.79	-1.17	-0.08	0.72	0.81	0.54	0.00
<i>D9S61</i>	-6.27	-2.28	-0.34	0.67	1.27	1.17	0.71	0.00
<i>ABL</i>	-19.34	-8.11	-3.69	-1.95	-0.55	-0.07	0.04	0.00
<i>D9S66</i>	-22.67	-5.40	-1.88	-0.55	0.37	0.53	0.36	0.00
<i>D9S67</i>	-7.91	-0.84	0.35	0.68	0.74	0.55	0.29	0.00

Lod scores were calculated at different recombination fractions. *Endoglin* lies between *D9S315* and *D9S61*.

**TABLE 4.7 TWO POINT LOD SCORES FOR MARKERS ADJACENT
TO *ALK1* IN FAMILY P**

Locus	Recombination fraction (θ)							
	0.00	0.01	0.05	0.10	0.20	0.30	0.40	0.50
<i>D12S331</i>	-19.30	-7.40	-3.80	-2.24	-0.90	-0.34	-0.09	0.00
<i>D12S85</i>	-3.77	-3.46	-1.75	-0.98	-0.35	-0.10	-0.01	0.00
<i>D12S347</i>	-24.37	-7.21	-3.55	-2.08	-0.85	-0.32	-0.08	0.00
<i>D12S368</i>	-20.77	-10.46	-5.79	-3.72	-1.80	-0.87	-0.34	0.00

Lod scores were calculated at different recombination fractions.

TABLE 4.8 CALCULATING ALLELE FREQUENCIES OF MARKER
***D9S315* IN FAMILY P**

Allele number	Total visualised on gel	Calculated allele frequency
1	1	0.017
2	9	0.155
3	14	0.242
4	10	0.172
5	1	0.017
6	8	0.138
7	4	0.069
8	11	0.190

Lod scores with altered allele frequencies reflecting those present in family P could then be calculated.

**TABLE 4.9 ANALYSIS OF LOD SCORES WITH EQUAL AND
ADJUSTED ALLELE FREQUENCIES IN FAMILY P**

Family	Locus	Allele freq.	Recombination fraction (θ)							
			0.00	0.01	0.05	0.10	0.20	0.30	0.40	0.50
P	<i>D9S315</i>	equal	-11.86	-3.79	-1.17	-0.08	0.72	0.81	0.54	0.00
P	<i>D9S315</i>	adjusted	-11.61	-3.54	-1.14	0.09	0.79	0.85	0.55	0.00
P	<i>D9S61</i>	equal	-6.27	-2.28	-0.34	0.67	1.27	1.17	0.71	0.00
P	<i>D9S61</i>	adjusted	-6.69	-2.62	-0.45	0.53	1.13	1.05	0.63	0.00

No significant difference was found between lod scores obtained using equal and adjusted allele frequencies.

**TABLE 4.10 TWO POINT LOD SCORES FOR MARKERS ADJACENT
TO *ALK1* IN FAMILY V AND FAMILY T**

Family V

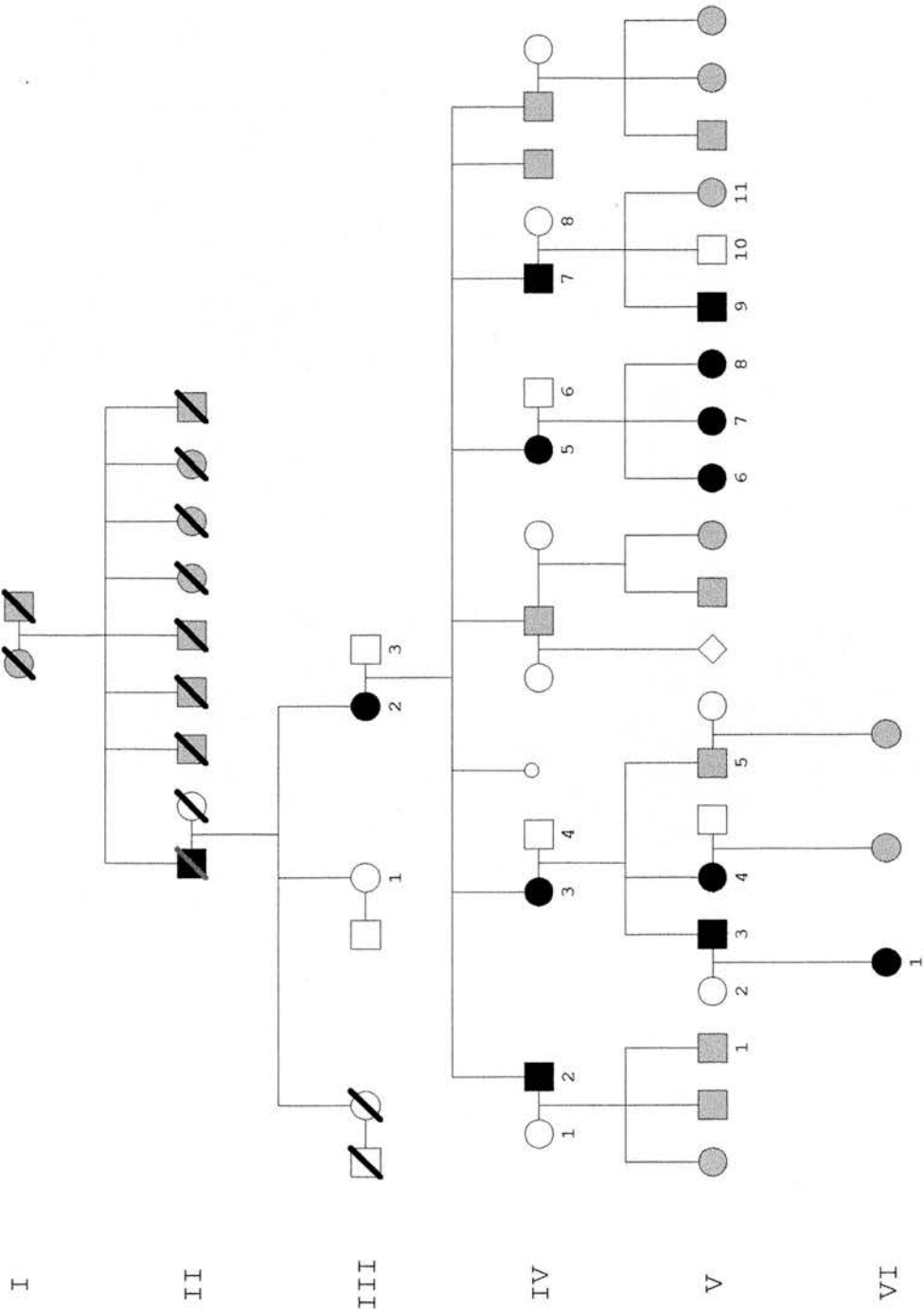
Locus	Recombination fraction (θ)							
	0.00	0.01	0.05	0.10	0.20	0.30	0.40	0.50
<i>D12S347</i>	-0.04	-0.02	0.03	0.65	0.07	0.05	0.01	0.00
<i>D12S368</i>	1.48	1.45	1.33	1.17	0.81	0.43	0.11	0.00

Family T

Locus	Recombination fraction (θ)							
	0.00	0.01	0.05	0.10	0.20	0.30	0.40	0.50
<i>D12S339</i>	-6.25	-1.82	-0.97	-0.58	-0.23	-0.08	-0.02	0.00
<i>D12S347</i>	0.49	0.47	0.43	0.37	0.24	0.12	0.03	0.00
<i>D12S368</i>	-5.98	-1.31	-0.51	-0.19	0.03	0.05	0.02	0.00

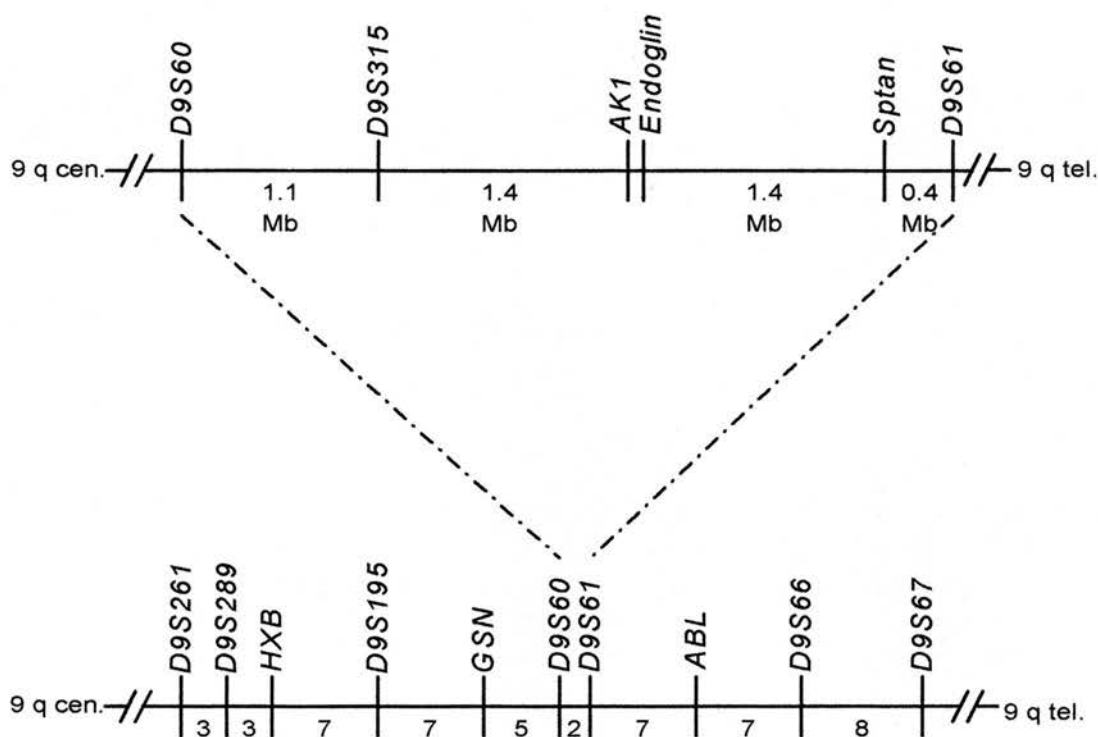
Lod scores were calculated at different recombination fractions.

FIGURE 4.1 FAMILY TREE OF FAMILY S



The numbered family members are the ones which were assessed clinically. Affected individuals are indicated by black symbols, unaffected by white symbols, and those with unknown diagnosis by grey symbols. Squares represent males, circles represent females, and a slashed symbol indicates a deceased individual.

FIGURE 4.2 THE POSITION OF *ENDOGLIN* AND ITS FLANKING MARKERS



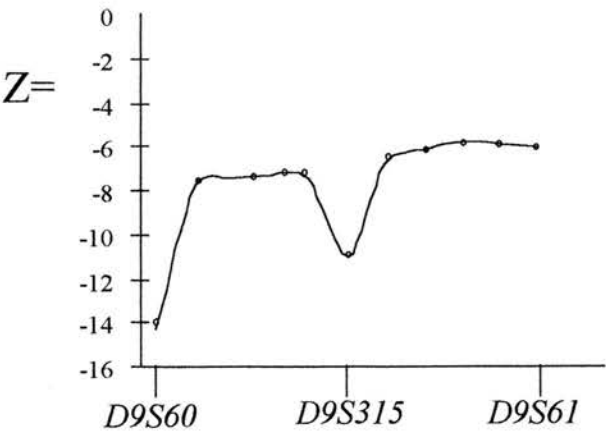
The position of *endoglin* relative to flanking markers. *Endoglin* lies in a 2 centiMorgan (cM) interval between *D9S60* and *D9S61*. The relative positions of *endoglin* and *AK1* have not been resolved. Distances in the upper part of the diagram, within the 2 cM region, are physical distances in megabases (Mb). Distances in the lower part of the diagram, showing the position of the *D9S60-D9S61* interval relative to other markers are genetic distances in centiMorgans. They are represented by the numbers below the horizontal line.

FIGURE 4.3 TWO POINT AND MULTIPOINT LOD SCORES
SPANNING THE *ENDOGLIN* LOCUS IN FAMILY S

A: Two point lod scores

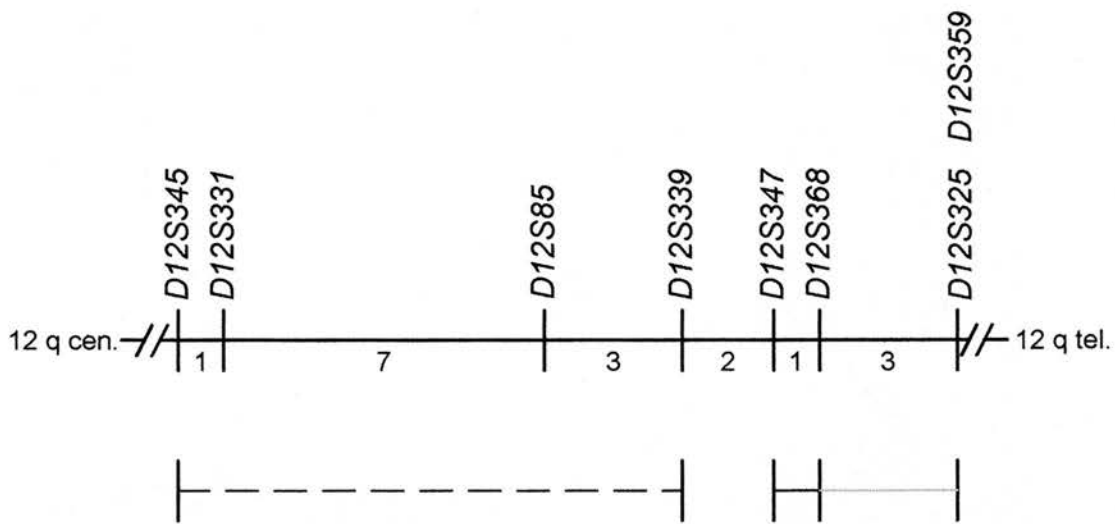
Locus	Recombination fraction (θ)							
	0.00	0.01	0.05	0.10	0.20	0.30	0.40	0.50
<i>D9S60</i>	-16.67	-8.99	-5.38	-3.52	-1.72	-0.79	-0.26	0.00
<i>D9S315</i>	-16.07	-7.67	-4.93	-3.21	-1.57	-0.72	-0.25	0.00
<i>D9S61</i>	-10.33	-5.58	-2.88	-1.74	-0.74	-0.31	-0.11	0.00

B: Multipoint lod scores



Data obtained using GENEHUNTER via HGMP.

FIGURE 4.4 THE POSITION OF *ALK1* AND ITS FLANKING MARKERS



Position of *ALK1* relative to flanking markers. The numbers below the horizontal line are the genetic distances in centiMorgans. The probable *ALK1* site is indicated by a solid black line, with possible extension indicated by a solid grey bar¹³⁹. The broken line indicates one of the original intervals¹³⁸ which has now been revised¹³⁹.

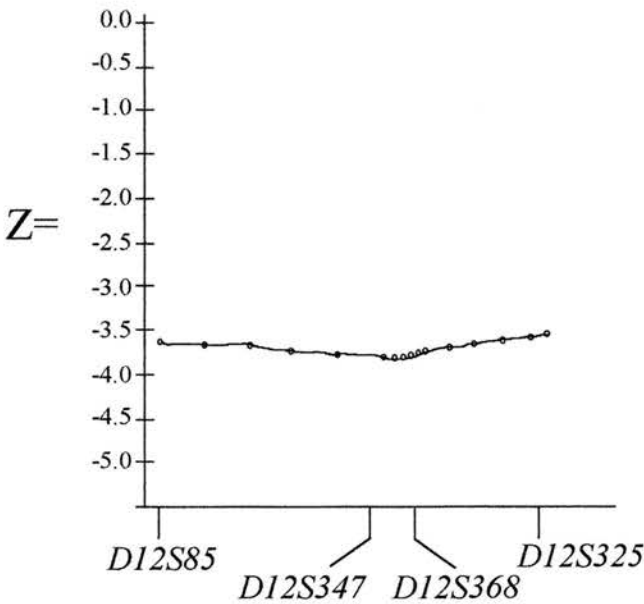
FIGURE 4.5 TWO POINT AND MULTIPOINT LOD SCORES
SPANNING THE *ALK1* LOCUS IN FAMILY S

A: Two point lod scores

Locus	Recombination fraction (θ)							
	0.00	0.01	0.05	0.10	0.20	0.30	0.40	0.50
<i>D12S331</i>	-4.81	-4.76	-3.39	-2.26	-1.17	-0.59	-0.23	0.00
<i>D12S85</i>	-4.67	-4.64	-4.57	-4.19	-2.39	-1.25	-0.50	0.00
<i>D12S347</i>	-9.07	-5.14	-4.60	-3.89	-2.16	-1.14	-0.47	0.00
<i>D12S368</i>	-5.27	-1.67	-0.97	-0.68	-0.39	-0.22	-0.10	0.00
<i>D12S325</i>	-8.78	-4.84	-4.34	-4.00	-2.38	-1.29	-0.55	0.00

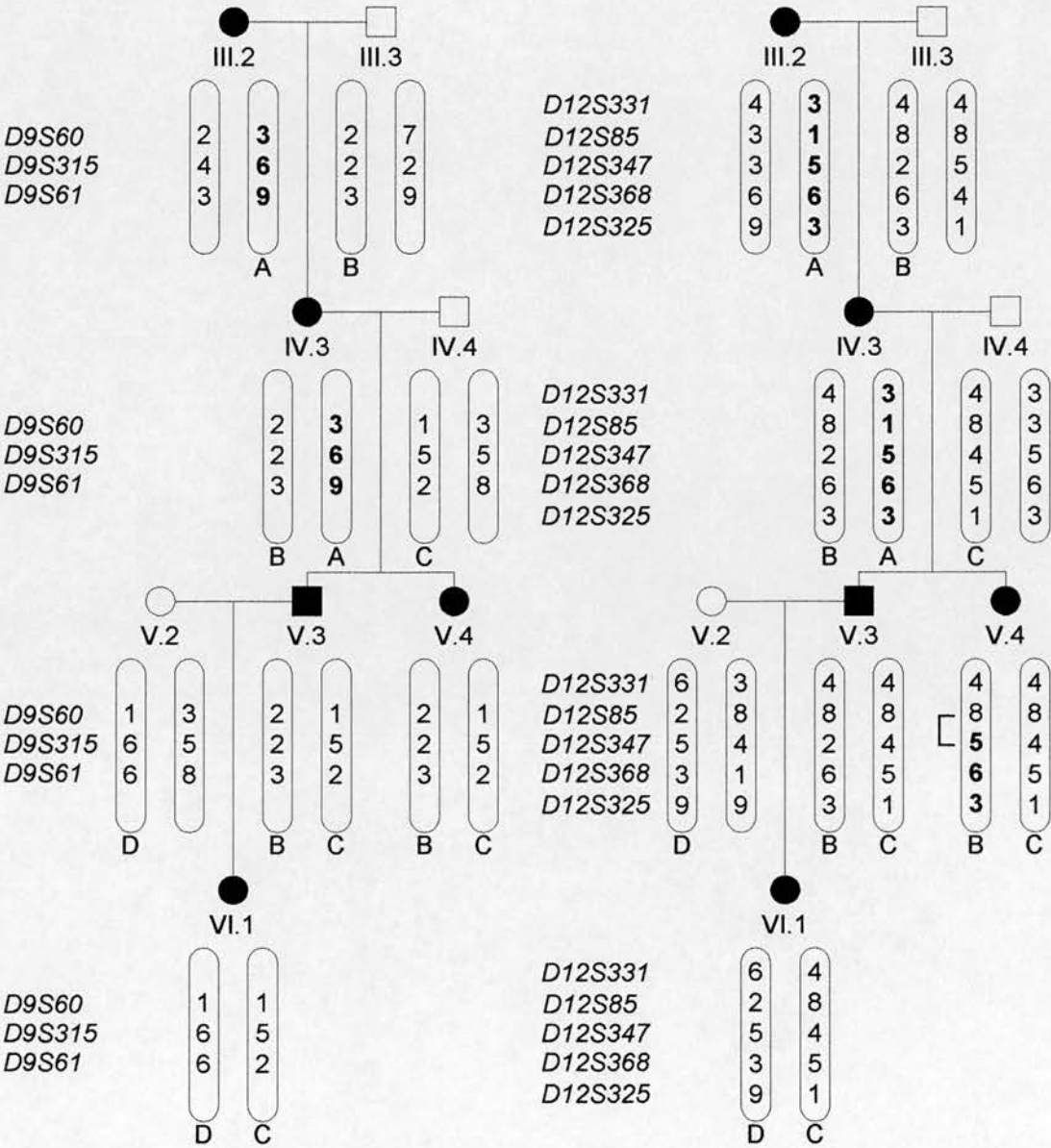
Lod scores were calculated at different recombination fractions.

B: Multipoint lod scores



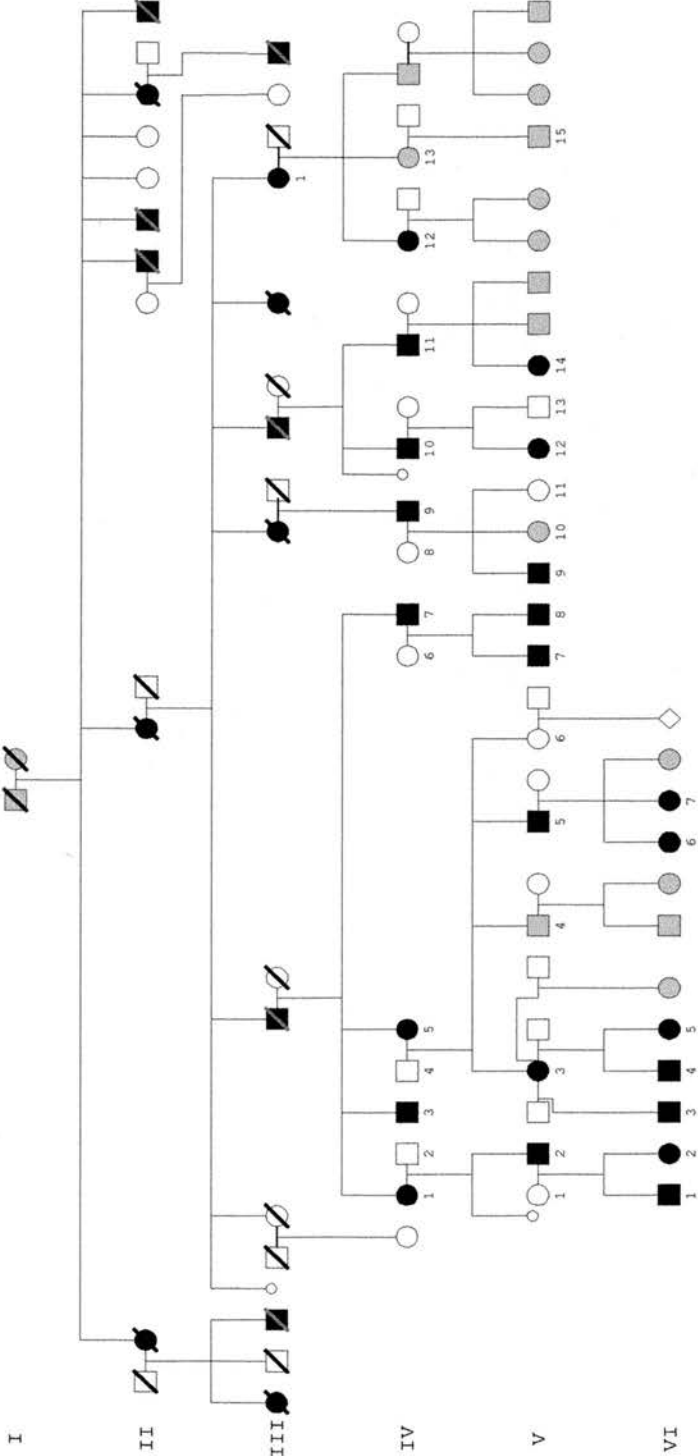
Data obtained using GENEHUNTER via HGMP.

FIGURE 4.6 **LIMITED HAPLOTYPE ANALYSES OF**
CHROMOSOME 9 AND 12 MARKERS IN FAMILY S



The analyses are limited to a section of the family S pedigree which includes the two PAVM individuals (V.3 and V.4). The pedigree shown on the left displays markers spanning *endoglin*. In affected individuals V.3, V.4, and VI.1 all haplotypes (B, C, and D) are derived from unaffected spouses. The pedigree on the right shows markers spanning *ALK1*. In affected individuals V.3, V.4, and VI.1 all haplotypes (B, C, and D) are derived from unaffected spouses. There is a recombination event in V.4 between *D12S85* and *D12S347*.

FIGURE 4.7 **FAMILY TREE OF FAMILY P**



The numbered family members are the ones which were assessed clinically. Affected individuals are indicated by black symbols, unaffected by white symbols, and those with unknown diagnosis by grey symbols. Squares represent males, circles represent females, and a slashed symbol indicates a deceased individual.

FIGURE 4.8

ILLUSTRATION OF ALLELE INHERITANCE OF
CHROMOSOME 9 MARKERS IN FAMILY P

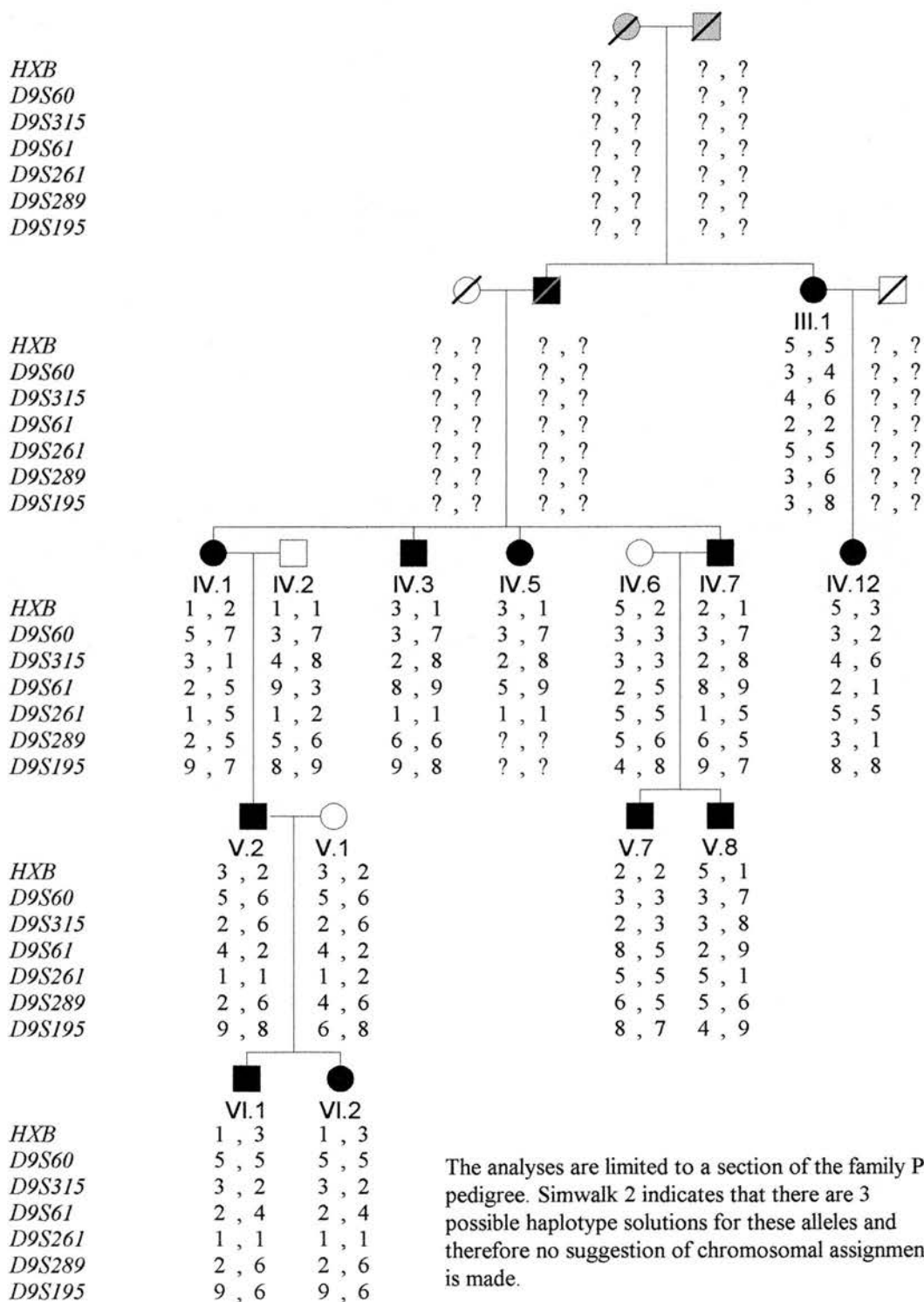
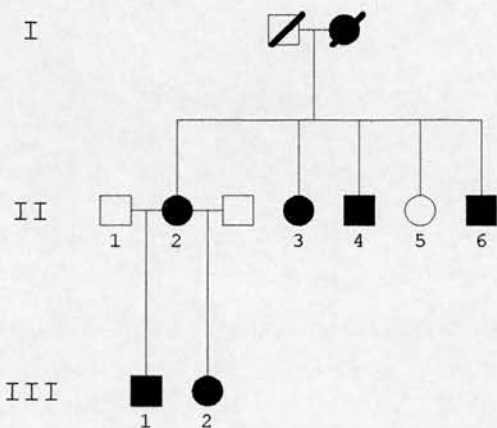
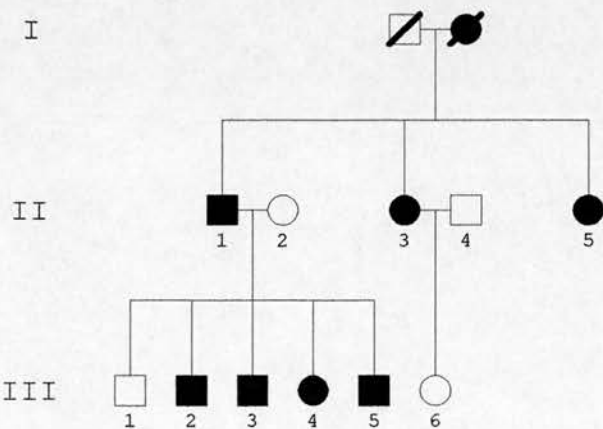


FIGURE 4.9 FAMILY TREE OF FAMILY V AND FAMILY T

Family V



Family T



The numbered family members are the ones which were assessed clinically. Affected individuals are indicated by black symbols, unaffected by white symbols, and those with unknown diagnosis by grey symbols. Squares represent males, circles represent females, and a slashed symbol indicates a deceased individual. Individual III.1 from family V and individual III.4 from family T had proven PAVMs.

5. DISCUSSION

This work has shown that a two base pair CA deletion polymorphism, with allele frequency of 9% in the UK population, occurs just over 2 kb upstream of the *endoglin* promoter. No polymorphisms or mutations have been found in the *endoglin* promoter in the HHT families studied.

The work has also shown that an HHT family is unlinked to the two known genes, *endoglin* and *ALK1*, indicating that at least a third gene exists for HHT. Moreover it has shown that PAVMs occur in the family linked to this third gene and also occur in an *ALK1*-linked family.

A number of issues arise from this work as detailed below.

5.1. CLINICAL ISSUES

5.1.1. PREGNANCY

A number of important clinical features have been highlighted by the work on families S and P. Family S member V.4 demonstrated that pregnancy can precipitate the onset of HHT manifestations. It is unknown whether family P member V.3's PAVMs were present prior to her pregnancy but family S member V.4 demonstrates that pregnancy can precipitate symptomatic PAVMs. The difficulty in diagnosing PAVMs during pregnancy, due to hypoxaemia being masked, is shown by individual V.3 from family P who had saturations of 98% supine, 97% erect during pregnancy but of 94/95% supine, 93/92% erect after delivery. The tendency for pregnancy to exacerbate HHT manifestations has been demonstrated by the deterioration of a telangiectatic lesion of the thumb in family P member V.3 during pregnancy, which improved after delivery of her child. It is not clear the extent to which haemodynamic changes or the very high oestrogen state^{48 131 132} contribute to the exacerbating effect of pregnancy on HHT manifestations.

5.1.1.1. PAVM RESCREENING

It can be seen from the discussion on pregnancy that rescreening for PAVMs is critical following pregnancy. Family P members V.5, VI.1, and VI.3 in whom pregnancy has obviously not occurred, have also shown that PAVMs have possibly developed since screening 6 years previously. This emphasises the need for rescreening all HHT patients. Family S member V.3 demonstrates that rescreening for PAVMs should be performed in patients with previously embolised PAVMs since his had possibly redeveloped.

The importance for rescreening for the HHT phenotype is highlighted by the work as 5 new diagnoses were made in individuals assessed 6 years previously. As age-related penetrance occurs in this disease the probability of becoming affected decreases with age, but should never be completely overlooked.

5.1.2. MICROSCOPIC PAVMS

Individual V.4 from family S demonstrates a case of microscopic PAVMs. Their clinical significance is under debate both in terms of their possible interference with gas exchange, and the possibility of paradoxical embolism occurring through them⁹⁵. As they may allow gas exchange, and so may lead to underestimation of the shunt using the 100% inspired oxygen breathing method¹¹⁹, technetium scanning may be more appropriate in these cases. Individual V.4's oximetry would suggest otherwise. Her frequent dizzy spells, possibly indicative of TIAs, would suggest that paradoxical embolism can occur.

If these microscopic PAVMs are clinically relevant then they must be detected. Nearly 50% of patients in whom contrast echocardiography is suggestive of PAVMs will not have PAVMs demonstrated at pulmonary angiography¹¹¹. While it is not clear whether these patients have microscopic PAVMs or whether they represent a subgroup of normal appearances⁹⁵, there is an increasingly prevalent view that contrast echocardiography may represent an appropriate first line screen for PAVMs¹¹¹. V.4's case suggests that contrast echocardiography may not detect all PAVMs however, since the contrast echocardiography was negative on the same day that her right to left shunt was quantified by technetium scanning as 7.8%. It would

support the view that contrast echocardiography should be used in conjunction with other investigations.

It could be argued that it is pointless detecting these microscopic PAVMs, even if they are clinically relevant, since embolisation is usually not possible. It is important however if they are clinically relevant for these patients to take prophylactic antibiotics before undergoing procedures that could cause bacteraemia⁹⁵. It is also vital to detect these patients so that they can be studied further, thus increasing our knowledge of the clinical implications and course of their disease.

5.2. MOLECULAR ISSUES

5.2.1. THE IMPORTANCE OF FINDING A POLYMORPHISM CLOSE TO THE *ENDOGLIN* PROMOTER

As discussed in Section 1.5.6 a polymorphism is useful as a tool for association and linkage studies. The polymorphism found just over 2 kb upstream of the *endoglin* promoter was detected prior to the publication of the human SNP map²²¹. Although of less importance now, at the time this finding was useful because the polymorphism has a reasonably high allele frequency, and there were then only seven known polymorphisms^{37 235 236} in the *endoglin* gene, the closest to the promoter being in exon 1.

5.2.1.1. SHOULD THE *ENDOGLIN* PROMOTER BE SCREENED FOR HHT MUTATIONS?

Since no mutations or even polymorphisms were seen in the *endoglin* promoter region in the eight HHT families screened in this work, the question arises as to whether or not the promoter should be screened in other HHT families. It remains possible that promoter mutations could account for disease in other *endoglin*-linked families given that they have null alleles. Moreover polymorphisms in the promoter regions of many other genes lead to disease. For instance in the promoter region of the *Factor VII* gene encoding a key component of the coagulation cascade, polymorphisms have been found which influence the binding properties of nuclear proteins and thereby alter the transcription rate of the gene²³⁷. A further polymorphism in this promoter has been found to alter factor VII levels and to be associated with differing risk of myocardial infarction amongst patients with an equal

extent of coronary atherosclerosis, independent of other risk factors²³⁸. Similarly polymorphisms in the protein C promoter region lead to different protein C levels and to different thrombotic risks²³⁹. These promoter polymorphisms are not just restricted to members of the coagulation cascade. Interestingly a polymorphism in the TNF α promoter causing increased transcription of the gene has been found to lead to an increased risk of cerebral malaria independent of HLA class I and II variation²⁴⁰.

It therefore seems appropriate to continue screening the *endoglin* promoter region for polymorphisms in other HHT families which may harbour polymorphisms that are functional and possibly disease-causing.

5.2.1.2. FUNCTIONAL EFFECTS OF ANY PROMOTER POLYMORPHISM FOUND

Although it is tempting on finding a sequence variation in the coding or regulatory regions of the gene to assume this will lead to a functional difference, scientific rigour is needed to confirm the hypothesis. Where present in an inherited disease, cosegregation of the variant with the disease in question needs to be demonstrated. Lack of cosegregation of the promoter polymorphism described previously (see Section 3.4.2) significantly reduced its likely functional importance.

Direct functional studies can be performed relatively easily by ligating either the wild-type promoter, or the promoter containing the promoter polymorphism, into separate reporter vectors containing a promoterless gene, where transcription of the gene can be assayed in transiently transfected cells. (This approach was used for

analysis of luciferase gene expression controlled by the wild-type *endoglin* promoter, with luciferase activity measured using a luminometer¹³³).

Similarly, confirmation of a disruption of function in a gene in which a coding sequence variation is found supports the hypothesis that the sequence variation is significant, and may shed light on the molecular mechanism of a particular pathological process. cDNA containing the wild-type or variant sequence can be ligated into a vector carrying a strong eukaryotic cell promoter, and either biochemical studies of the encoded proteins, or studies of downstream effects on cell function, may be compared. This approach was successfully used to confirm the haploinsufficient disease mechanism for *endoglin* in HHT¹⁴⁹, and disruption of normal *endoglin* function in ongoing studies in my supervisor's laboratory²⁴¹.

5.2.1.3. OTHER REGIONS TO BE SCREENED

Although the promoter region and entire coding sequence of *endoglin* has been screened without detecting the disease-causing mutation in four *endoglin*-linked HHT families, further regions of the gene may contain the disease-causing mutation. Introns (the first intron in particular) and 3' untranslated regions tend to contain regulatory elements for many genes, and a testable hypothesis would be that haploinsufficiency of *endoglin* could also result due to mutations in these regulatory elements.

5.2.2. THE IMPORTANCE OF IDENTIFYING THE THIRD GENE

As described in the introduction chapter, *endoglin* and *ALK1* are both involved in the regulation of response to ligands of the TGF β superfamily. It seems reasonable to speculate that the gene mutated in family S, the family unlinked to *endoglin* and

ALK1, may encode another component of the TGF β signalling complex or a downstream effector. Its identification should significantly enhance our understanding of signalling by TGF β family members which, although increasing, still has many areas yet to be elucidated.

As discussed in the introduction chapter, the TGF β superfamily is implicated in many different pathological processes including vascular development, homeostasis, and repair, and cellular proliferation^{185 186 188 189}. These are important in a wide range of disease processes including atherosclerosis, oncogenesis, pulmonary hypertension and fibrosis in addition to HHT. There are therefore compelling clinical as well as scientific reasons why understanding the mechanism by which TGF- β family members act is so important.

The confirmation by this work that at least a third HHT gene exists has precipitated work to identify the gene. Using genomic DNA of affected and unaffected members of family S, the family unlinked to *endoglin* and *ALK1*, PCR-based analysis of polymorphic markers in the vicinity of candidate genes selected from the TGF β superfamily has been performed. No linkage was found to these genes²⁴² and so a genome-wide search using fluorescently-labelled polymorphic markers with an average 10cM resolution has been commenced. Over a third of the genome has been excluded to date. Further details of the findings of this part of the analysis are not available to me. Any regions in which lod scores are suggestive of linkage in the initial screen using markers at 10cM intervals will be saturated by analysis of further polymorphisms in the vicinity. Once a lod score greater than 3.0 is obtained,

analyses of GenBank (<http://www.ncbi.nlm.nih.gov>) and Ensembl (<http://www.ensembl.org>) databases will be performed to identify potential candidate genes in the region. This will lead to mutational screening of the suspected gene in this family by Southern blotting and sequencing of PCR products from genomic DNA and cDNA. The genomic DNA analyses will be substantially assisted by the Ensembl database of genomic DNA sequence which can be scrutinised using short DNA sequences analogous to Blast searches of the GenBank databases.

5.2.2.1. ELUCIDATION OF THE BIOCHEMICAL ROLE

In this study, the work did not reach the stage of identification of the third gene and therefore elucidation of its biochemical role. However, as mentioned previously, my work has precipitated work to identify the third gene and thereafter its biochemical role. Analyses are first likely to include clarification as to whether or not the mutant protein is required for the pathogenesis of the disease, by assessment of production of mRNA or protein by the mutant allele. A low level of rogue mRNA transcription is predicted in all cells²⁴³ and therefore the cDNA of the gene in question is likely to be amplified by reverse transcriptase PCR (RT-PCR) amplification of mRNA isolated from peripheral blood leukocytes or EBV-transformed lymphoblastoid lines of an affected individual³⁷. Protein analysis requires significantly higher levels of transcription and production of a stable protein. This is best performed in overexpression systems by transiently transfecting eukaryotic cells such as Chinese Hamster Ovary (CHO) cells with the normal (wild type) gene under investigation and with various mutant constructs of the gene. The effects of the mutant constructs can then be studied by a number of biochemical methods including

immunoprecipitation with radiolabelled ligands or radiolabelled antibodies to the encoded protein, Western blot analysis, or flow cytometry if antibodies to the new gene are available (alternative metabolic labelling methods are also available if monoclonal antibodies are not available)^{146 149}. Functional studies can be performed assessing for general physiological effects such as alterations of cellular proliferation or migration^{189 244}, or effects on specific genes by assessment of the cellular concentrations of particular mRNA species by Northern blots¹⁸⁹, or altered function of a promoter thought likely to be affected by the signalling from the wild-type protein (analysed using a luciferase reporter plasmid known to be affected by the protein under investigation²⁴¹). Ultimately, many studies of gene function progress to assessment of transgenic animals in which the gene has been deleted by homologous recombination (leading to null or “knock-out” mice), or in which a gene is overexpressed, a technique applicable to all species as it is not dependent on the availability of pluripotent embryonal stem (ES) cells, currently available for mice only^{150 245}.

5.2.3. PHENOTYPE/ GENOTYPE CORRELATIONS

PAVMs have previously been seen predominantly in *endoglin*-linked families^{37 143 144}. This led to the suggestion that a correlation could be made between pulmonary involvement and genotype^{141 142}. This work shows that PAVMs also occur in families linked to *ALK1* and at least a third gene. There is therefore no absolute association between pulmonary involvement and genotype. This has significant implications for clinical PAVM screening, the importance of which has been discussed previously. From these data it can be recommended that all HHT patients

are screened clinically for PAVMs. These data are also important for mutational screening programmes. Patients with pulmonary involvement, whose underlying disease gene is unknown, should not be limited to *endoglin* screening but should have linkage analysis performed first if family size permits.

Since PAVMs cannot be used as an indicator of genotype, the question arises as to whether there are any phenotypic differences in HHT manifestation between family S, the family unlinked to *endoglin* and *ALK1*, and other HHT families linked to these genes. There do not appear to be phenotypic differences. As with other families this pedigree showed a range of severity of disease. The PAVM clinical presentation and response to embolisation of V.3 reflect the majority of PAVM patients who present symptomatically. The apparent precipitation of V.4's symptoms and PAVMs by pregnancy is well recognised¹⁰⁶. The incidence of cerebral and hepatic involvement in this family is unknown as asymptomatic screening for these complications is not performed routinely in the UK. The geographical and fish tongue seen in IV.3 could be coincidental or could be associated with HHT, although it has not been described previously. It was however also seen in members of family P which has not been proven to be linked to a third gene. At the moment, therefore, possible phenotypic indicators of underlying genotype remain elusive.

5.3. VASCULAR BIOLOGY ISSUES

5.3.1. WHY ARE SOME VESSELS ABNORMAL AND NOT OTHERS?

A fundamental question remains which is why do abnormal vascular structures, in particular PAVMs, develop? *Endoglin* and *ALK1* are both involved in the TGF β superfamily which has a role in vasculogenesis and angiogenesis so one might expect low levels of *endoglin* or *ALK1* in HHT affected individuals to lead to abnormal vessel morphology. However AVMs and apparently normal vessels in affected individuals have similar levels of *endoglin* (and presumably *ALK1*) so there must be additional factors leading to AVM formation¹⁶⁴. Altered blood flow and/or hormonal changes could initiate the process in vessels already susceptible due to decreased *endoglin/ALK1* levels⁹⁵. Vessel remodelling could then be exacerbated through a lack of the homeostatic mechanisms normally orchestrated by TGF β I¹⁸⁸. It could be that other genetic factors, as yet undetermined, could influence which vessels develop AVMs¹⁶⁴.

5.4. CONCLUDING REMARKS

Although much is now known about *endoglin*, and to a lesser extent *ALK1*, regarding specific gene mutations and the biochemical actions of the proteins they encode, a clear understanding of the pathophysiology of HHT has yet to be achieved.

Identification of a third gene and elucidation of its encoded protein should increase our knowledge of the actions of the TGF β superfamily, but may not immediately improve our understanding of why disruption of these actions can lead to abnormal vascular structures. It is vital that research continues in this fascinating area to improve our knowledge of the pathophysiology of both HHT and also the many other diseases in which the TGF β superfamily is clearly involved.

5.5. FURTHER WORK

The results of my work highlight a number of areas requiring further study, several of which have already commenced:

5.5.1. CLINICAL

- From a clinical point of view, further study of patients with microscopic PAVMs is important to determine their clinical significance, diagnosis, and possible treatment. The risks versus the benefits of screening and treating asymptomatic cerebral AVMs needs to be determined so that a decision can be made on screening for these CAVMs²⁴⁶. Continuing clinical assessment of HHT families is vital if we are to identify any phenotypic indicators of genotype.

5.5.2. MOLECULAR

- The *endoglin* promoter region should be screened in further *endoglin*-linked HHT families: (I have been asked by a Canadian group, currently screening the endoglin gene for mutations, to supply them with the primer sequences and screening methods used in this study.)
- Regions of the first intron (unlikely to be in its entirety as the intron is over 12kb long) and the area 3' to exon 14 should be screened in the families studied in this work, and further families in which the *endoglin* mutation is elusive.
- The confirmation of the existence of at least a third HHT gene necessitates the identification of the gene. Once the gene is identified this will precipitate further work to elucidate its biochemical role. Since this thesis was first written this work has commenced, first by analysing for linkage to candidate genes from the

TGF β superfamily and then by a genome-wide search using conventional linkage techniques²⁴².

5.5.3. VASCULAR BIOLOGY

- More generally, further work is required to identify the *endoglin* and *ALK1* ligand and the particular TGF β superfamily ligands involved in HHT. The mechanism by which the TGF β superfamily leads to the abnormal vascular structures of HHT, and indeed to the other many pathological processes with which it is involved, must be elucidated if molecular therapies are going to be offered in the future.

BIBLIOGRAPHY

1. Sutton HG. Epistaxis as an indication of impaired nutrition, and of degeneration of the vascular system. *Medical Mirror* 1864:769-81.
2. Babington BG. Hereditary epistaxis. *Lancet* 1865;2:362-3.
3. Legg W. A case of haemophilia complicated with multiple naevi. *Lancet* 1876;2:856.
4. Chiari O. Erfahrungen auf dem gebiete der hals und nasenkrankheiten. Vienna, 1887:60 et seq.
5. Rendu HJ. Epistaxis repetees chez un sujet porteur de petits angiomes cutanes et muqueux. *Gazette des Hopitaux* 1896:1322-3.
6. Osler W. On a family form of recurring epistaxis, associated with multiple telangiectases of the skin and mucous membranes. *Johns Hopkins Hospital Bulletin* 1901;12:333-7.
7. Hawthorne CO. Recurring epistaxis with multiple telangiectases of the skin. *Lancet* 1906;1:90.
8. Kelly AB. Multiple telangiectases of the skin and mucous membranes of the nose and mouth. *Glasgow Medical Journal* 1906;65(6).
9. Kelly AB. Multiple telangiectases of the skin and mucous membranes of the nose and throat. *Proceedings of the Royal Society of Medicine* 1908;1(5):45.
10. Waggett EB. A case of multiple telangiectases. *Proceedings of the Royal Society of Medicine, Laryngological Section* 1908;1(6).
11. Weber FP. Multiple hereditary developmental angiomas (telangiectases) of the skin and mucous membranes associated with recurring haemorrhages. *Lancet* 1907;2:160-2.
12. Hanes FM. Multiple hereditary telangiectases causing hemorrhage (hereditary hemorrhagic telangiectasia). *Johns Hopkins Hospital Bulletin* 1909;20:63-73.
13. Malpighus MchY, J.). De Pulmonibus. Cited in *Proceedings of the Royal Society of Medicine* 1661 (cited 1929):1-10.
14. Churton. Multiple aneurysm of pulmonary artery. *British Medical Journal* 1897;1:1223.
15. Smith HL, Horton BT. Arteriovenous fistula of the lung associated with polycythaemia vera: report of a case in which the diagnosis was made clinically. *American Heart Journal* 1939;18:589-92.

16. Rundles RW. Hemorrhagic telangiectasia with pulmonary artery aneurysm: case report. *American Journal of Medical Science* 1945;210:76-81.
17. Renshaw JF. Multiple hemorrhagic telangiectasis with special reference to gastroscopic appearance. *Cleveland Clinic Quarterly* 1936;6:226-30.
18. Whitaker W. Cavernous haemangioma of the lung. *Thorax* 1947;2:58-64.
19. Schwartz SO, Armstrong BE. Familial hereditary hemorrhagic telangiectasia in the negro. *New England Journal of Medicine* 1948;239(12):434-5.
20. Smith JL, Lineback MI. Hereditary hemorrhagic telangiectasia: nine cases in one negro family, with special reference to hepatic lesions. *American Journal of Medicine* 1954:41-8.
21. Guillén B, Guízar J, de la Cruz J, Salamanca F. Hereditary hemorrhagic telangiectasia: report of 15 affected cases in a Mexican family. *Clinical Genetics* 1991;39(3):214-8.
22. Cheah JS, Ghosh MB, Tan SK. Hereditary haemorrhagic telangiectasia (Rendu-Osler-Weber disease) in a Chinese family. *Singapore Medical Journal* 1970;11(3):168-72.
23. Chauhan MS, Ahuja JM, Tewari SC, Jayaswal R, Dandona PK. Osler-Rendu-Weber disease presenting as recurrent haemoptysis. *Indian Journal of Chest Disease and Allied Sciences* 1989;31(3):227-32.
24. Panich V, Wasi P, Na-Nakorn S. Hereditary hemorrhagic telangiectasia. The first report in Thailand. *Journal of the Medical Association of Thailand* 1972;55(3):195-206.
25. Chen K, Oda K, Iwasaki T, Yorifuji S, Tomomatsu T. [A case of intrahepatic arteriovenous fistula in Rendu-Osler-Weber disease]. *Nippon Naika Gakkai Zasshi* 1972;61(10):1305-13.
26. Plauchu H, Bideau A. Epidemiologie et constitution d'un registre de population a propos d'une concentration géographique d'une maladie héréditaire rare. *Population* 1984;4-5:765-86.
27. Vase P, Grove O. Gastrointestinal lesions in hereditary hemorrhagic telangiectasia. *Gastroenterology* 1986;91(5):1079-83.
28. Jessurun GA, Kamphuis DJ, van der Zande FH, Nossent JC. Cerebral arteriovenous malformations in The Netherlands Antilles. High prevalence of hereditary hemorrhagic telangiectasia-related single and multiple cerebral arteriovenous malformations. *Clinical Neurology and Neurosurgery* 1993;95(3):193-8.

29. Guttmacher AE, McKinnon WC, Upton MD. Hereditary hemorrhagic telangiectasia: a disorder in search of the genetics community [letter]. *American Journal of Medical Genetics* 1994;52(2):252-3.
30. Porteous ME, Burn J, Proctor SJ. Hereditary haemorrhagic telangiectasia: a clinical analysis. *Journal of Medical Genetics* 1992;29(8):527-30.
31. Kjeldsen AD, Vase P, Green A. Hereditary haemorrhagic telangiectasia: a population-based study of prevalence and mortality in Danish patients. *Journal of Internal Medicine* 1999;245(1):31-9.
32. Garland HG, Anning ST. Hereditary haemorrhagic telangiectasia: a genetic and bibliographical study. *British Journal of Dermatology and Syphilis* 1950;62:289-309.
33. McKusick VA. Mendelian Inheritance in Man; catalogs of autosomal dominant, autosomal recessive, and X-linked phenotypes. *Mendelian Inheritance in Man*. Baltimore: Johns Hopkins University Press, 1990:902-4.
34. Plauchu H, de Chadarevian JP, Bideau A, Robert JM. Age-related clinical profile of hereditary hemorrhagic telangiectasia in an epidemiologically recruited population. *American Journal of Medical Genetics* 1989;32(3):291-7.
35. Shovlin CL, Hughes JM, Tuddenham EG, Temperley I, Perembelon YF, Scott J, et al. A gene for hereditary haemorrhagic telangiectasia maps to chromosome 9q3. *Nature Genetics* 1994;6(2):205-9.
36. Shovlin CL, Hughes JM. Hereditary hemorrhagic telangiectasia [letter; comment]. *New England Journal of Medicine* 1996;334(5):330-1; discussion 331-2.
37. Shovlin CL, Hughes JM, Scott J, Seidman CE, Seidman JG. Characterization of endoglin and identification of novel mutations in hereditary hemorrhagic telangiectasia. *American Journal of Human Genetics* 1997;61(1):68-79.
38. Snyder LH. Is the homozygous form of multiple telangiectasia lethal? *The Journal of Laboratory and Clinical Medicine* 1944;29:1211-16.
39. Muller JY, Michailov T, Izrael V, Bernard J. [Hereditary haemorrhagic telangiectasia in a large Saharan family. 87 cases in the same family (author's transl)]. *Nouvelle Presse Medicale* 1978;7(20):1723-5.
40. Anabtawi IA, Ellison RG, Ellison LT. Pulmonary arteriovenous aneurysms and fistulas. *Annals of Thoracic Surgery* 1965;1:277-85.
41. Bosher L, Blake A, Byrd B. An analysis of the pathologic anatomy of pulmonary arteriovenous aneurysms with particular reference to the applicability of local excision. *Surgery* 1959;45:91-104.

42. Hashimoto K, Pritzker MS. Hereditary hemorrhagic telangiectasia- an electron microscopic study. *Oral Surgery* 1972;34(5):751-68.
43. Menefee MG, Flessa HC, Glueck HI, Hogg SP. Hereditary hemorrhagic telangiectasia (Osler-Weber-Rendu disease). An electron microscopic study of the vascular lesions before and after therapy with hormones. *Archives of Otolaryngology* 1975;101(4):246-51.
44. Braverman IM, Keh A, Jacobson BS. Ultrastructure and three-dimensional organization of the telangiectases of hereditary hemorrhagic telangiectasia. *Journal of Investigative Dermatology* 1990;95(4):422-7.
45. Assar OS, Friedman CM, White RIJ. The natural history of epistaxis in hereditary hemorrhagic telangiectasia. *Laryngoscope* 1991;101(9):977-80.
46. Stecker RH, Lake CF. Hereditary hemorrhagic telangiectasia; review of 102 cases and presentation of an innovation to septodermoplasty. *Archives of Otolaryngology* 1965;82(5):522-6.
47. McCaffrey TV, Kern EB, Lake CF. Management of epistaxis in hereditary hemorrhagic telangiectasia. Review of 80 cases. *Archives of Otolaryngology* 1977;103(11):627-30.
48. Harrison DF. Use of estrogen in treatment of familial hemorrhagic telangiectasia. *Laryngoscope* 1982;92(3):314-20.
49. Bird RM, Hammarsten JF, Marshall RA, Robinson RR. A study of hereditary hemorrhagic telangiectasia. *New England Journal of Medicine* 1957;257(3):105-9.
50. Parkin JL, Dixon JA. Laser photocoagulation in hereditary hemorrhagic telangiectasia. *Otolaryngology - Head and Neck Surgery* 1981;89(2):204-8.
51. Shapshay SM, Oliver P. Treatment of hereditary hemorrhagic telangiectasia by Nd-YAG laser photocoagulation. *Laryngoscope* 1984;94(12 Pt 1):1554-6.
52. Saunders WH. Septal dermoplasty for hereditary telangiectasia and other conditions. *Otolaryngology Clinic of North America* 1973;6(3):745-55.
53. Ulsø C, Vase P, Stoksted P. Long-term results of dermatoplasty in the treatment of hereditary haemorrhagic telangiectasia. *Journal of Laryngology and Otology* 1983;97(3):223-6.
54. Vase P. Estrogen treatment of hereditary hemorrhagic telangiectasia. A double-blind controlled clinical trial. *Acta Medica Scandinavica* 1981;209(5):393-6.
55. Vase P, Lorentzen M. Histological findings following oestrogen treatment of hereditary haemorrhagic telangiectasia. A controlled double-blind investigation. *Journal of Laryngology and Otology* 1983;97(5):427-9.

56. Richtsmeier W, Weaver G, Streck W, Jacobson H, Dewell R, Olson J. Estrogen and progesterone receptors in hereditary hemorrhagic telangiectasia. *Otolaryngology - Head and Neck Surgery* 1984;92(5):564-70.
57. Merland JJ, Melki JP, Chiras J, Riche MC, Hadjean E. Place of embolization in the treatment of severe epistaxis. *Laryngoscope* 1980;90(10 Pt 1):1694-704.
58. Parnes LS, Heeneman H, Vinuela F. Percutaneous embolization for control of nasal blood circulation. *Laryngoscope* 1987;97(11):1312-5.
59. Elden L, Montanera W, Terbrugge K, Willinsky R, Lasjaunias P, Charles D. Angiographic embolization for the treatment of epistaxis: a review of 108 cases. *Otolaryngology - Head and Neck Surgery* 1994;111(1):44-50.
60. Hodgson CH, Burchell HB, Good CA, Clagett OT. Hereditary hemorrhagic telangiectasia and pulmonary arteriovenous fistula. *New England Journal of Medicine* 1959;261(13):625-36.
61. Hodgson CH, Kaye RL. Pulmonary arteriovenous fistula and hereditary hemorrhagic telangiectasia: a review and report of 35 cases of fistula. *Diseases of the Chest* 1963;43(5):449-55.
62. Brant AM, Schachat AP, White RI. Ocular manifestations in hereditary hemorrhagic telangiectasia (Rendu-Osler-Weber disease). *American Journal of Ophthalmology* 1989;107(6):642-6.
63. Morelli JG, Huff JC, Weston WL. Treatment of congenital telangiectatic vascular malformations with the pulsed-dye laser (585 nm). *Pediatrics* 1993;92(4):603-6.
64. Reilly PJ, Nostrant TT. Clinical manifestations of hereditary hemorrhagic telangiectasia. *American Journal of Gastroenterology* 1984;79(5):363-7.
65. Gostout CJ, Bowyer BA, Ahlquist DA, Viggiano TR, Balm RK. Mucosal vascular malformations of the gastrointestinal tract: clinical observations and results of endoscopic neodymium: yttrium-aluminum- garnet laser therapy. *Mayo Clinic Proceedings* 1988;63(10):993-1003.
66. Halpern M, Turner AF, Citron BP. Hereditary hemorrhagic telangiectasia. An angiographic study of abdominal visceral angiodysplasias associated with gastrointestinal hemorrhage. *Radiology* 1968;90(6):1143-9.
67. Rutgeerts P, Van Gompel F, Geboes K, Vantrappen G, Broeckaert L, Coremans G. Long term results of treatment of vascular malformations of the gastrointestinal tract by neodymium Yag laser photocoagulation. *Gut* 1985;26(6):586-93.

68. Sargeant IR, Loizou LA, Rampton D, Tulloch M, Bown SG. Laser ablation of upper gastrointestinal vascular ectasias: long term results. *Gut* 1993;34(4):470-5.
69. Van Cutsem E. Georges Brohee Prize. Oestrogen-progesterone, a new therapy of bleeding gastrointestinal vascular malformations. *Acta Gastroenterologica Belgica* 1993;56(1):2-10.
70. Korzenik JR, Topazian MD, White R. Treatment of bleeding in hereditary hemorrhagic telangiectasia with aminocaproic acid [letter; comment]. *New England Journal of Medicine* 1994;331(18):1236.
71. Saba HI, Morelli GA, Logrono LA. Brief report: treatment of bleeding in hereditary hemorrhagic telangiectasia with aminocaproic acid [see comments]. *New England Journal of Medicine* 1994;330(25):1789-90.
72. Buscarini E, Buscarini L, Danesino C, Piantanida M, Civardi G, Quaretti P, et al. Hepatic vascular malformations in hereditary hemorrhagic telangiectasia: Doppler sonographic screening in a large family. *Journal of Hepatology* 1997;26(1):111-8.
73. Caselitz M, Wagner S, Chavan A, Gebel M, Bleck JS, Wu A, et al. Clinical outcome of transfemoral embolisation in patients with arteriovenous malformations of the liver in hereditary haemorrhagic telangiectasia (Weber-Rendu-Osler disease). *Gut* 1998;42(1):123-6.
74. Bernard G, Mion F, Henry L, Plauchu H, Paliard P. Hepatic involvement in hereditary hemorrhagic telangiectasia: clinical, radiological, and hemodynamic studies of 11 cases. *Gastroenterology* 1993;105(2):482-7.
75. Fagel WJ, Perlberger R, Kauffmann RH. Portosystemic encephalopathy in hereditary hemorrhagic telangiectasia. *American Journal of Medicine* 1988;85(6):858-60.
76. Zentler-Munro PL, Howard ER, Karani J, Williams R. Variceal haemorrhage in hereditary haemorrhagic telangiectasia [see comments]. *Gut* 1989;30(9):1293-7.
77. Bjørø K, Schrumpf E, Elgjo K, Kolmanskog F. Monstrous ascites in hereditary haemorrhagic telangiectasia. *Scandinavian Journal of Gastroenterology* 1995;30(1):92-4.
78. Martini GA. The liver in hereditary haemorrhagic teleangiectasia: an inborn error of vascular structure with multiple manifestations: a reappraisal. *Gut* 1978;19(6):531-7.
79. Cooney T, Sweeney EC, Coll R, Greally M. 'Pseudocirrhosis' in hereditary haemorrhagic telangiectasia. *Journal of Clinical Pathology* 1977;30(12):1134-41.

80. Daly JJ, Schiller AL. The liver in hereditary hemorrhagic telangiectasia (Osler-Weber-Rendu disease). *American Journal of Medicine* 1976;60(5):723-6.
81. Ralls PW, Johnson MB, Radin DR, Lee KP, Boswell WD. Hereditary hemorrhagic telangiectasia: findings in the liver with color Doppler sonography. *American Journal of Roentgenology* 1992;159(1):59-61.
82. Haitjema T, Westermann CJ, Overtom TT, Timmer R, Disch F, Mauser H, et al. Hereditary hemorrhagic telangiectasia (Osler-Weber-Rendu disease): new insights in pathogenesis, complications, and treatment [see comments]. *Archives of Internal Medicine* 1996;156(7):714-9.
83. Bauer T, Britton P, Lomas D, Wight DG, Friend PJ, Alexander GJ. Liver transplantation for hepatic arteriovenous malformation in hereditary haemorrhagic telangiectasia. *Journal of Hepatology* 1995;22(5):586-90.
84. Miller FJ, Jr., Whiting JH, Korzenik JR, White RI. Caution with use of hepatic embolization in the treatment of hereditary hemorrhagic telangiectasia [letter; comment]. *Radiology* 1999;213(3):928-30.
85. Kitchens CS, Lottenberg R. Chronic painless hematuria and urethral bleeding as the presenting manifestations of Osler-Weber-Rendu disease. *Journal of Urology* 1976;116(5):681-2.
86. Haitjema T, Disch F, Overtom TT, Westermann CJ, Lammers JW. Screening family members of patients with hereditary hemorrhagic telangiectasia. *American Journal of Medicine* 1995;99(5):519-24.
87. Román G, Fisher M, Perl DP, Poser CM. Neurological manifestations of hereditary hemorrhagic telangiectasia (Rendu-Osler-Weber disease): report of 2 cases and review of the literature. *Annals of Neurology* 1978;4(2):130-44.
88. Graf CJ, Perret GE, Torner JC. Bleeding from cerebral arteriovenous malformations as part of their natural history. *Journal of Neurosurgery* 1983;58(3):331-7.
89. Turjman F, Massoud TF, Vinuela F, Sayre JW, Guglielmi G, Duckwiler G. Correlation of the angioarchitectural features of cerebral arteriovenous malformations with clinical presentation of hemorrhage. *Neurosurgery* 1995;37(5):856-60; discussion 860-2.
90. Willinsky RA, Lasjaunias P, Terbrugge K, Burrows P. Multiple cerebral arteriovenous malformations (AVMs). Review of our experience from 203 patients with cerebral vascular lesions. *Neuroradiology* 1990;32(3):207-10.
91. ter Berg JW, Dippel DW, Habbema JD, Westermann CJ, Tulleken CA, Willemsse J. Unruptured intracranial arteriovenous malformations with hereditary haemorrhagic telangiectasia. Neurosurgical treatment or not? *Acta Neurochirurgica* 1993;121(1-2):34-42.

92. Wallace GMF, Jackson JE, Hughes JMB, Shovlin CL. The risk of cerebral haemorrhage in families with HHT and pulmonary AVMs. *Thorax* 1999;54 (suppl 3):32.
93. Rigamonti D, Hadley MN, Drayer BP, Johnson PC, Hoenig-Rigamonti K, Knight JT, et al. Cerebral cavernous malformations. Incidence and familial occurrence. *New England Journal of Medicine* 1988;319(6):343-7.
94. Pikus HJ, Beach ML, Harbaugh RE. Microsurgical treatment of arteriovenous malformations: analysis and comparison with stereotactic radiosurgery [see comments]. *Journal of Neurosurgery* 1998;88(4):641-6.
95. Shovlin CL, Letarte M. Hereditary haemorrhagic telangiectasia and pulmonary arteriovenous malformations: issues in clinical management and review of pathogenic mechanisms. *Thorax* 1999;54(8):714-29.
96. Vase P, Holm M, Arendrup H. Pulmonary arteriovenous fistulas in hereditary hemorrhagic telangiectasia. *Acta Medica Scandinavica* 1985;218(1):105-9.
97. White RI, Jr., Lynch-Nyhan A, Terry P, Buescher PC, Farmlett EJ, Charnas L, et al. Pulmonary arteriovenous malformations: techniques and long-term outcome of embolotherapy. *Radiology* 1988;169(3):663-9.
98. Swanson KL, Prakash UB, Stanson AW. Pulmonary arteriovenous fistulas: Mayo Clinic experience, 1982-1997. *Mayo Clinic Proceedings* 1999;74(7):671-80.
99. Dines DE, Arms RA, Bernatz PE, Gomes MR. Pulmonary arteriovenous fistulas. *Mayo Clinic Proceedings* 1974;49(7):460-5.
100. Dines DE, Seward JB, Bernatz PE. Pulmonary arteriovenous fistulas. *Mayo Clinic Proceedings* 1983;58(3):176-81.
101. Sluiter-Eringa H, Orie NG, Sluiter HJ. Pulmonary arteriovenous fistula. Diagnosis and prognosis in noncomplainant patients. *American Review of Respiratory Disease* 1969;100(2):177-88.
102. Burke CM, Safai C, Nelson DP, Raffin TA. Pulmonary arteriovenous malformations: a critical update. *American Review of Respiratory Disease* 1986;134(2):334-9.
103. Gomes MR, Bernatz PE, Dines DE. Pulmonary arteriovenous fistulas. *Annals of Thoracic Surgery* 1969;7(6):582-93.
104. Dalton ML, Jr., Goodwin FC, Bronwell AW, Rutledge R. Intrapleural rupture of pulmonary arteriovenous aneurysm. Report of a case. *Diseases of the Chest* 1967;52(1):97-100.

105. Ference BA, Shannon TM, White RIJ, Zawin M, Burdge CM. Life-threatening pulmonary hemorrhage with pulmonary arteriovenous malformations and hereditary hemorrhagic telangiectasia. *Chest* 1994;106(5):1387-90.
106. Shovlin CL, Winstock AR, Peters AM, Jackson JE, Hughes JM. Medical complications of pregnancy in hereditary haemorrhagic telangiectasia. *Quarterly Journal of Medicine* 1995;88(12):879-87.
107. Haitjema TJ, Overtom TT, Westermann CJ, Lammers JW. Embolisation of pulmonary arteriovenous malformations: results and follow up in 32 patients [see comments]. *Thorax* 1995;50(7):719-23.
108. Whyte MK, Hughes JM, Jackson JE, Peters AM, Hempleman SC, Moore DP, et al. Cardiopulmonary response to exercise in patients with intrapulmonary vascular shunts. *Journal of Applied Physiology* 1993;75(1):321-8.
109. White RI, Jr. Pulmonary arteriovenous malformations: how do we diagnose them and why is it important to do so? [editorial; comment]. *Radiology* 1992;182(3):633-5.
110. Kjeldsen AD, Vase P, Oxhøj H. Hereditary hemorrhagic telangiectasia [letter; comment]. *New England Journal of Medicine* 1996;334(5):331-2.
111. Kjeldsen AD, Oxhøj H, Andersen PE, Elle B, Jacobsen JP, Vase P. Pulmonary arteriovenous malformations: screening procedures and pulmonary angiography in patients with hereditary hemorrhagic telangiectasia. *Chest* 1999;116(2):432-9.
112. Remy J, Remy-Jardin M, Wattinne L, Deffontaines C. Pulmonary arteriovenous malformations: evaluation with CT of the chest before and after treatment [see comments]. *Radiology* 1992;182(3):809-16.
113. Nickerson BG, Sarkisian C, Tremper K. Bias and precision of pulse oximeters and arterial oximeters. *Chest* 1988;93(3):515-7.
114. Thompson RD, Jackson J, Peters AM, Doré CJ, Hughes JM. Sensitivity and specificity of radioisotope right-left shunt measurements and pulse oximetry for the early detection of pulmonary arteriovenous malformations. *Chest* 1999;115(1):109-13.
115. Ueki J, Hughes JM, Peters AM, Bellingan GJ, Mohammed MA, Dutton J, et al. Oxygen and 99mTc-MAA shunt estimations in patients with pulmonary arteriovenous malformations: effects of changes in posture and lung volume. *Thorax* 1994;49(4):327-31.
116. Andersen PE, Kjeldsen AD, Oxhøj H, Vase P, White RIJ. Embolotherapy for pulmonary arteriovenous malformations in patients with hereditary hemorrhagic telangiectasia (Rendu-Osler-Weber syndrome). *Acta Radiologica* 1998;39(6):723-6.

117. Guttmacher AE, Marchuk DA, White RI. Hereditary Hemorrhagic Telangiectasia (letter; comment). *New England Journal of Medicine* 1996;334(5):331-2.
118. Chilvers ER, Peters AM, George P, Hughes JM, Allison DJ. Quantification of right to left shunt through pulmonary arteriovenous malformations using ⁹⁹Tcm albumin microspheres. *Clinical Radiology* 1988;39(6):611-4.
119. Genovesi MG, Tierney DF, Taplin GV, Eisenberg H. An intravenous radionuclide method to evaluate hypoxemia caused by abnormal alveolar vessels. Limitation of conventional techniques. *American Review of Respiratory Disease* 1976;114(1):59-65.
120. Shub C, Tajik AJ, Seward JB, Dines DE. Detecting intrapulmonary right-to-left shunt with contrast echocardiography. Observations in a patient with diffuse pulmonary arteriovenous fistulas. *Mayo Clinic Proceedings* 1976;51(2):81-4.
121. Chilvers ER, Whyte MK, Jackson JE, Allison DJ, Hughes JM. Effect of percutaneous transcatheter embolization on pulmonary function, right-to-left shunt, and arterial oxygenation in patients with pulmonary arteriovenous malformations. *American Review of Respiratory Disease* 1990;142(2):420-5.
122. Dutton JA, Jackson JE, Hughes JM, Whyte MK, Peters AM, Ussov W, et al. Pulmonary arteriovenous malformations: results of treatment with coil embolization in 53 patients. *American Journal of Roentgenology* 1995;165(5):1119-25.
123. Haitjema T, ten Berg JM, Overtom TT, Ernst JM, Westermann CJ. Unusual complications after embolization of a pulmonary arteriovenous malformation. *Chest* 1996;109(5):1401-4.
124. Faughnan ME, Lui YW, Wirth JA, Pugash RA, Redelmeier DA, Hyland RH, et al. Diffuse pulmonary arteriovenous malformations: characteristics and prognosis. *Chest* 2000 Jan;117(1):31-8.
125. Watanabe N, Munakata Y, Ogiwara M, Miyatake M, Nakagawa F, Hirayama J. A case of pulmonary arteriovenous malformation in a patient with brain abscess successfully treated with video-assisted thoracoscopic resection. *Chest* 1995;108(6):1724-7.
126. Reynaud-Gaubert M, Thomas P, Gaubert JY, Pietri P, Garbe L, Giudicelli R, et al. Pulmonary arteriovenous malformations: lung transplantation as a therapeutic option. *European Respiratory Journal* 1999;14(6):1425-8.
127. White RI, Jr. Case 16-1990; brain abscess and pulmonary arteriovenous malformations [letter]. *New England Journal of Medicine* 1991;324(20):1439-40.

128. Livneh A, Langevitz P, Morag B, Catania A, Pras M. Functionally reversible hepatic arteriovenous fistulas during pregnancy in patients with hereditary hemorrhagic telangiectasia. *Southern Medical Journal* 1988;81(8):1047-9.
129. Moore LG, Reeves JT. Pregnancy blunts pulmonary vascular reactivity in dogs. *American Journal of Physiology* 1980;239(3):H297-301.
130. Guyton AC. *Textbook of Medical Physiology*. seventh ed: Saunders, 1986.
131. Koch HJ, Escher GC, Lewis SJ. Hormonal management of hereditary hemorrhagic telangiectasia. *Journal of the American Medical Association (JAMA)* 1952;149:1376-80.
132. Rowley PT, Kurnick J, Cheville R. Hereditary haemorrhagic telangiectasia: aggravation by oral contraceptives? *Lancet* 1970;1(7644):474-5.
133. Rius C, Smith JD, Almendro N, Langa C, Botella LM, Marchuk DA, et al. Cloning of the promoter region of human endoglin, the target gene for hereditary hemorrhagic telangiectasia type 1. *Blood* 1998;92(12):4677-90.
134. British Thoracic Society SoCC. Suspected acute pulmonary embolism: a practical approach. *Thorax* 1997;52 Suppl 4:S1-24.
135. McDonald MT, Papenberg KA, Ghosh S, Glatfelter AA, Biesecker BB, Helmbold EA, et al. A disease locus for hereditary haemorrhagic telangiectasia maps to chromosome 9q33-34. *Nature Genetics* 1994;6(2):197-204.
136. McAllister KA, Grogg KM, Johnson DW, Gallione CJ, Baldwin MA, Jackson CE, et al. Endoglin, a TGF-beta binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. *Nature Genetics* 1994;8(4):345-51.
137. Vincent P, Plauchu H, Hazan J, Fauré S, Weissenbach J, Godet J. A third locus for hereditary haemorrhagic telangiectasia maps to chromosome 12q [published erratum appears in Hum Mol Genet 1995 Jul;4(7):1243]. *Human Molecular Genetics* 1995;4(5):945-9.
138. Johnson DW, Berg JN, Gallione CJ, McAllister KA, Warner JP, Helmbold EA, et al. A second locus for hereditary hemorrhagic telangiectasia maps to chromosome 12. *Genome Research* 1995;5(1):21-8.
139. Johnson DW, Berg JN, Baldwin MA, Gallione CJ, Marondel I, Yoon SJ, et al. Mutations in the activin receptor-like kinase 1 gene in hereditary haemorrhagic telangiectasia type 2. *Nature Genetics* 1996;13(2):189-95.
140. Piantanida M, Buscarini E, Dellavecchia C, Minelli A, Rossi A, Buscarini L, et al. Hereditary haemorrhagic telangiectasia with extensive liver involvement is not caused by either HHT1 or HHT2. *Journal of Medical Genetics* 1996;33(6):441-3.

141. McAllister KA, Lennon F, Bowles-Biesecker B, McKinnon WC, Helmbold EA, Markel DS, et al. Genetic heterogeneity in hereditary haemorrhagic telangiectasia: possible correlation with clinical phenotype. *Journal of Medical Genetics* 1994;31(12):927-32.
142. Porteous ME, Curtis A, Williams O, Marchuk D, Bhattacharya SS, Burn J. Genetic heterogeneity in hereditary haemorrhagic telangiectasia. *Journal of Medical Genetics* 1994;31(12):925-6.
143. Heutink P, Haitjema T, Breedveld GJ, Janssen B, Sandkuijl LA, Bontekoe CJ, et al. Linkage of hereditary haemorrhagic telangiectasia to chromosome 9q34 and evidence for locus heterogeneity. *Journal of Medical Genetics* 1994;31(12):933-6.
144. Berg JN, Guttmacher AE, Marchuk DA, Porteous ME. Clinical heterogeneity in hereditary haemorrhagic telangiectasia: are pulmonary arteriovenous malformations more common in families linked to endoglin? *Journal of Medical Genetics* 1996;33(3):256-7.
145. McAllister KA, Baldwin MA, Thukkani AK, Gallione CJ, Berg JN, Porteous ME, et al. Six novel mutations in the endoglin gene in hereditary hemorrhagic telangiectasia type 1 suggest a dominant-negative effect of receptor function. *Human Molecular Genetics* 1995;4(10):1983-5.
146. Pece N, Vera S, Cymerman U, White RIJ, Wrana JL, Letarte M. Mutant endoglin in hereditary hemorrhagic telangiectasia type 1 is transiently expressed intracellularly and is not a dominant negative. *Journal of Clinical Investigation* 1997;100(10):2568-79.
147. Yamaguchi H, Azuma H, Shigekiyo T, Inoue H, Saito S. A novel missense mutation in the endoglin gene in hereditary hemorrhagic telangiectasia. *Thrombosis and Haemostasis* 1997;77(2):243-7.
148. Gallione CJ, Klaus DJ, Yeh EY, Stenzel TT, Xue Y, Anthony KB, et al. Mutation and expression analysis of the endoglin gene in hereditary hemorrhagic telangiectasia reveals null alleles. *Human Mutation* 1998;11(4):286-94.
149. Paquet M-E, Pece-Barbara N, Vera S, Cymerman U, Karabegovic A, Shovlin C, et al. Analysis of several endoglin mutants reveals no endogenous mature or secreted protein capable of interfering with normal endoglin function, 2001.
150. Bourdeau A, Dumont DJ, Letarte M. A murine model of hereditary hemorrhagic telangiectasia [see comments]. *Journal of Clinical Investigation* 1999;104(10):1343-51.

151. Li DY, Sorensen LK, Brooke BS, Urness LD, Davis EC, Taylor DG, et al. Defective angiogenesis in mice lacking endoglin. *Science* 1999;284(5419):1534-7.
152. Gougos A, Letarte M. Identification of a human endothelial cell antigen with monoclonal antibody 44G4 produced against a pre-B leukemic cell line. *Journal of Immunology* 1988;141(6):1925-33.
153. St-Jacques S, Forte M, Lye SJ, Letarte M. Localization of endoglin, a transforming growth factor-beta binding protein, and of CD44 and integrins in placenta during the first trimester of pregnancy. *Biology of Reproduction* 1994;51(3):405-13.
154. St-Jacques S, Cymerman U, Pece N, Letarte M. Molecular characterization and in situ localization of murine endoglin reveal that it is a transforming growth factor-beta binding protein of endothelial and stromal cells. *Endocrinology* 1994;134(6):2645-57.
155. Bellón T, Corbí A, Lastres P, Calés C, Cebrián M, Vera S, et al. Identification and expression of two forms of the human transforming growth factor-beta-binding protein endoglin with distinct cytoplasmic regions. *European Journal of Immunology* 1993;23(9):2340-5.
156. Adam PJ, Clesham GJ, Weissberg PL. Expression of endoglin mRNA and protein in human vascular smooth muscle cells. *Biochemical and Biophysical Research Communications* 1998;247(1):33-7.
157. Qu R, Silver MM, Letarte M. Distribution of endoglin in early human development reveals high levels on endocardial cushion tissue mesenchyme during valve formation. *Cell and Tissue Research* 1998;292(2):333-43.
158. Buhring HJ, Muller CA, Letarte M, Gougos A, Saalmuller A, van Agthoven AJ, et al. Endoglin is expressed on a subpopulation of immature erythroid cells of normal human bone marrow. *Leukemia* 1991;5(10):841-7.
159. Lastres P, Bellon T, Cabañas C, Sanchez-Madrid F, Acevedo A, Gougos A, et al. Regulated expression on human macrophages of endoglin, an Arg-Gly-Asp-containing surface antigen. *European Journal of Immunology* 1992;22(2):393-7.
160. Wang JM, Kumar S, Pye D, van Agthoven AJ, Krupinski J, Hunter RD. A monoclonal antibody detects heterogeneity in vascular endothelium of tumours and normal tissues. *International Journal of Cancer* 1993;54(3):363-70.
161. Rulo HF, Westphal JR, van de Kerkhof PC, de Waal RM, van Vlijmen IM, Ruiter DJ. Expression of endoglin in psoriatic involved and uninvolved skin. *Journal of Dermatological Science* 1995;10(2):103-9.

162. Wang JM, Kumar S, van Agthoven A, Kumar P, Pye D, Hunter RD. Irradiation induces up-regulation of E9 protein (CD105) in human vascular endothelial cells. *International Journal of Cancer* 1995;62(6):791-6.
163. Guerrero-Esteo M, Lastres P, Letamendia A, Perez-Alvarez MJ, Langa C, Lopez LA, et al. Endoglin overexpression modulates cellular morphology, migration, and adhesion of mouse fibroblasts. *European Journal of Cell Biology* 1999;78(9):614-23.
164. Bourdeau A, Cymerman U, Paquet ME, Meschino W, McKinnon WC, Guttmacher AE, et al. Endoglin expression is reduced in normal vessels but still detectable in arteriovenous malformations of patients with hereditary hemorrhagic telangiectasia type 1. *American Journal of Pathology* 2000;156(3):911-23.
165. Pugh BF, Tjian R. Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* 1990;61(7):1187-97.
166. Kim SJ, Glick A, Sporn MB, Roberts AB. Characterization of the promoter region of the human transforming growth factor-beta 1 gene. *Journal of Biological Chemistry* 1989;264(1):402-8.
167. Geiser AG, Kim SJ, Roberts AB, Sporn MB. Characterization of the mouse transforming growth factor-beta 1 promoter and activation by the Ha-ras oncogene. *Molecular and Cellular Biology* 1991;11(1):84-92.
168. Schmitt J, Hotten G, Jenkins NA, Gilbert DJ, Copeland NG, Pohl J, et al. Structure, chromosomal localization, and expression analysis of the mouse inhibin/activin beta C (Inhbc) gene. *Genomics* 1996;32(3):358-66.
169. Bae HW, Geiser AG, Kim DH, Chung MT, Burmester JK, Sporn MB, et al. Characterization of the promoter region of the human transforming growth factor-beta type II receptor gene. *Journal of Biological Chemistry* 1995;270(49):29460-8.
170. Humphries DE, Bloom BB, Fine A, Goldstein RH. Structure and expression of the promoter for the human type II transforming growth factor-beta receptor. *Biochemical and Biophysical Research Communications* 1994;203(2):1020-7.
171. Resnick N, Gimbrone MA, Jr. Hemodynamic forces are complex regulators of endothelial gene expression. *Faseb Journal* 1995;9(10):874-82.
172. Geiser AG, Busam KJ, Kim SJ, Lafyatis R, MA OR, Webbink R, et al. Regulation of the transforming growth factor-beta 1 and -beta 3 promoters by transcription factor Sp1. *Gene* 1993;129(2):223-8.
173. Ungefroren H, Krull NB. Transcriptional regulation of the human biglycan gene. *Journal of Biological Chemistry* 1996;271(26):15787-95.

174. Ji C, Casinghino S, McCarthy TL, Centrella M. Multiple and essential Sp1 binding sites in the promoter for transforming growth factor-beta type I receptor. *Journal of Biological Chemistry* 1997;272(34):21260-7.
175. Choi SG, Yi Y, Kim YS, Kato M, Chang J, Chung HW, et al. A novel ets-related transcription factor, ERT/ESX/ESE-1, regulates expression of the transforming growth factor-beta type II receptor. *Journal of Biological Chemistry* 1998;273(1):110-7.
176. Yingling JM, Datto MB, Wong C, Frederick JP, Liberati NT, Wang XF. Tumor suppressor Smad4 is a transforming growth factor beta-inducible DNA binding protein. *Molecular and Cellular Biology* 1997;17(12):7019-28.
177. Stroschein SL, Wang W, Luo K. Cooperative binding of Smad proteins to two adjacent DNA elements in the plasminogen activator inhibitor-1 promoter mediates transforming growth factor beta-induced smad-dependent transcriptional activation. *Journal of Biological Chemistry* 1999;274(14):9431-41.
178. Berg JN, Gallione CJ, Stenzel TT, Johnson DW, Allen WP, Schwartz CE, et al. The activin receptor-like kinase 1 gene: genomic structure and mutations in hereditary hemorrhagic telangiectasia type 2. *American Journal of Human Genetics* 1997;61(1):60-7.
179. Abdalla SA, Pece-Barbara N, Vera S, Tapia E, Paez E, Bernabeu C, et al. Analysis of ALK-1 and endoglin in newborns from families with hereditary hemorrhagic telangiectasia type 2. *Human Molecular Genetics* 2000;9(8):1227-37.
180. Oh SP, Seki T, Goss KA, Imamura T, Yi Y, Donahoe PK, et al. Activin receptor-like kinase 1 modulates transforming growth factor-beta 1 signaling in the regulation of angiogenesis. *Proceedings of the National Academy of Science, U S A (Proc Natl Acad Sci USA)* 2000;97(6):2626-31.
181. Attisano L, Cárcamo J, Ventura F, Weis FM, Massagué J, Wrana JL. Identification of human activin and TGF beta type I receptors that form heteromeric kinase complexes with type II receptors. *Cell* 1993;75(4):671-80.
182. ten Dijke P, Ichijo H, Franzén P, Schulz P, Saras J, Toyoshima H, et al. Activin receptor-like kinases: a novel subclass of cell-surface receptors with predicted serine/threonine kinase activity. *Oncogene* 1993;8(10):2879-87.
183. Wu X, Robinson CE, Fong HW, Crabtree JS, Rodriguez BR, Roe BA, et al. Cloning and characterization of the murine activin receptor like kinase-1 (ALK-1) homolog. *Biochemical and Biophysical Research Communications* 1995;216(1):78-83.

184. Massagué J. The transforming growth factor-beta family. *Annual Review of Cell Biology* 1990;6:597-641.
185. Dickson MC, Martin JS, Cousins FM, Kulkarni AB, Karlsson S, Akhurst RJ. Defective haematopoiesis and vasculogenesis in transforming growth factor-beta 1 knock out mice. *Development* 1995;121(6):1845-54.
186. Kaartinen V, Voncken JW, Shuler C, Warburton D, Bu D, Heisterkamp N, et al. Abnormal lung development and cleft palate in mice lacking TGF-beta 3 indicates defects of epithelial-mesenchymal interaction. *Nature Genetics* 1995;11(4):415-21.
187. Koh GY, Kim SJ, Klug MG, Park K, Soonpaa MH, Field LJ. Targeted expression of transforming growth factor-beta 1 in intracardiac grafts promotes vascular endothelial cell DNA synthesis. *Journal of Clinical Investigation* 1995;95(1):114-21.
188. Madri JA, L B, J.R. M. Modulation of vascular cell behaviour by transforming growth factors beta. *Molecular Reproduction and Development* 1992;32(2):121-6.
189. Lastres P, Letamendía A, Zhang H, Rius C, Almendro N, Raab U, et al. Endoglin modulates cellular responses to TGF-beta 1. *Journal of Cell Biology* 1996;133(5):1109-21.
190. Letterio JJ, Roberts AB. TGF-beta: a critical modulator of immune cell function. *Clinical Immunology and Immunopathology* 1997;84(3):244-50.
191. O'Kane S, Ferguson MW. Transforming growth factor betas and wound healing. *International Journal of Biochemical Cell Biology* 1997;29(1):63-78.
192. ten Dijke P, Franzén P, Yamashita H, Ichijo H, Heldin CH, Miyazono K. Serine/threonine kinase receptors. *Progress In Growth Factor Research* 1994;5(1):55-72.
193. Attisano L, Wrana JL. Mads and Smads in TGF beta signalling. *Current Opinion in Cell Biology* 1998;10(2):188-94.
194. Macias-Silva M, Hoodless PA, Tang SJ, Buchwald M, Wrana JL. Specific activation of Smad1 signaling pathways by the BMP7 type I receptor, ALK2. *Journal of Biological Chemistry* 1998;273(40):25628-36.
195. Wrana JL, Attisano L, Cárcamo J, Zentella A, Doody J, Laiho M, et al. TGF beta signals through a heteromeric protein kinase receptor complex. *Cell* 1992;71(6):1003-14.
196. Franzén P, Heldin CH, Miyazono K. The GS domain of the transforming growth factor-beta type I receptor is important in signal transduction. *Biochemical and Biophysical Research Communications* 1995;207(2):682-9.

197. Ebner R, Chen RH, Lawler S, Zioncheck T, Derynck R. Determination of type I receptor specificity by the type II receptors for TGF-beta or activin. *Science* 1993;262(5135):900-2.
198. ten Dijke P, Yamashita H, Ichijo H, Franzén P, Laiho M, Miyazono K, et al. Characterization of type I receptors for transforming growth factor-beta and activin. *Science* 1994;264(5155):101-4.
199. ten Dijke P, Yamashita H, Sampath TK, Reddi AH, Estevez M, Riddle DL, et al. Identification of type I receptors for osteogenic protein-1 and bone morphogenetic protein-4. *Journal of Biological Chemistry* 1994;269(25):16985-8.
200. Liu F, Ventura F, Doody J, Massagué J. Human type II receptor for bone morphogenetic proteins (BMPs): extension of the two-kinase receptor model to the BMPs. *Molecular and Cellular Biology* 1995;15(7):3479-86.
201. Nohno T, Ishikawa T, Saito T, Hosokawa K, Noji S, Wolsing DH, et al. Identification of a human type II receptor for bone morphogenetic protein-4 that forms differential heteromeric complexes with bone morphogenetic protein type I receptors. *Journal of Biological Chemistry* 1995;270(38):22522-6.
202. Macias-Silva M, Abdollah S, Hoodless PA, Pirone R, Attisano L, Wrana JL. MADR2 is a substrate of the TGFbeta receptor and its phosphorylation is required for nuclear accumulation and signaling. *Cell* 1996;87(7):1215-24.
203. Nakao A, Imamura T, Souchelnytskyi S, Kawabata M, Ishisaki A, Oeda E, et al. TGF-beta receptor-mediated signalling through Smad2, Smad3 and Smad4. *Embo Journal* 1997;16(17):5353-62.
204. Nakao A, Roijer E, Imamura T, Souchelnytskyi S, Stenman G, Heldin CH, et al. Identification of Smad2, a human Mad-related protein in the transforming growth factor beta signaling pathway. *Journal of Biological Chemistry* 1997;272(5):2896-900.
205. Hoodless PA, Haerry T, Abdollah S, Stapleton M, MB OC, Attisano L, et al. MADR1, a MAD-related protein that functions in BMP2 signaling pathways. *Cell* 1996;85(4):489-500.
206. Yamamoto N, Akiyama S, Katagiri T, Namiki M, Kurokawa T, Suda T. Smad1 and smad5 act downstream of intracellular signalings of BMP-2 that inhibits myogenic differentiation and induces osteoblast differentiation in C2C12 myoblasts. *Biochemical and Biophysical Research Communications* 1997;238(2):574-80.
207. Barbara NP, Wrana JL, Letarte M. Endoglin is an accessory protein that interacts with the signaling receptor complex of multiple members of the

- transforming growth factor-beta superfamily. *Journal of Biological Chemistry* 1999;274(2):584-94.
208. Gougos A, Letarte M. Primary structure of endoglin, an RGD-containing glycoprotein of human endothelial cells. *Journal of Biological Chemistry* 1990;265(15):8361-4.
 209. Cheifetz S, Bellón T, Calés C, Vera S, Bernabeu C, Massagué J, et al. Endoglin is a component of the transforming growth factor-beta receptor system in human endothelial cells. *Journal of Biological Chemistry* 1992;267(27):19027-30.
 210. Lopez-Casillas F, Cheifetz S, Doody J, Andres JL, Lane WS, Massague J. Structure and expression of the membrane proteoglycan betaglycan, a component of the TGF-beta receptor system. *Cell* 1991;67(4):785-95.
 211. López-Casillas F, Wrana JL, Massagué J. Betaglycan presents ligand to the TGF beta signaling receptor. *Cell* 1993;73(7):1435-44.
 212. López-Casillas F, Payne HM, Andres JL, Massagué J. Betaglycan can act as a dual modulator of TGF-beta access to signaling receptors: mapping of ligand binding and GAG attachment sites. *Journal of Cell Biology* 1994;124(4):557-68.
 213. Yamashita H, Ichijo H, Grimsby S, Morén A, ten Dijke P, Miyazono K. Endoglin forms a heteromeric complex with the signaling receptors for transforming growth factor-beta. *Journal of Biological Chemistry* 1994;269(3):1995-2001.
 214. Lastres P, Martín-Perez J, Langa C, Bernabéu C. Phosphorylation of the human-transforming-growth-factor-beta-binding protein endoglin. *Biochemical Journal* 1994;301 (Pt 3):765-8.
 215. Ott J. *Analysis of human genetic linkage*. 2 ed. Baltimore: Johns Hopkins university press, 1991.
 216. C.L.Shovlin Thesis [program], 1996.
 217. Shovlin CL, Guttmacher AE, Buscarini E, Faughnan ME, Hyland RH, Westermann CJ, et al. Diagnostic criteria for hereditary hemorrhagic telangiectasia (Rendu-Osler-Weber syndrome). *American Journal of Medical Genetics* 2000;91(1):66-7.
 218. Bell J, Haldane JBS. The linkage between genes for colour blindness and haemophilia. *Proceedings of the Royal Society, Series B* 1937;123:119-150.
 219. Renwick JH. Progress in mapping human autosomes. *British Medical Bulletin* 1969;25(1):65-73.

220. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 1980;32(3):314-31.
221. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 2001;409(6822):928-33.
222. Stoneking M. Single nucleotide polymorphisms. From the evolutionary past.. *Nature* 2001 Feb 15;409(6822):821-2.
223. Chakravarti A. To a future of genetic medicine. *Nature* 2001 Feb 15;409(6822):822-3.
224. Kunkel TA, Bebenek K. Recent studies of the fidelity of DNA synthesis. *Biochimica Biophysica Acta* 1988;951(1):1-15.
225. Kwiatkowski DJ, Dib C, Slaugenhaupt SA, Povey S, Gusella JF, Haines JL. An index marker map of chromosome 9 provides strong evidence for positive interference. *American Journal of Human Genetics* 1993;53(6):1279-88.
226. Keavney B. Genetic association studies in complex diseases. *Journal of Human Hypertension* 2000;14(6):361-7.
227. Collins FS, McKusick VA. Implications of the Human Genome Project for medical science. *Journal of the American Medical Association (JAMA)* 2001 Feb 7;285(5):540-4.
228. Ryder REJ, Mir MA, Freeman EA. *An aid to the MRCP short cases*. 2 ed. Cambridge: Blackwell Science, 1999.
229. Shovlin CL, Guttmacher AE, Buscarini E, Faughnan M, Hyland R, Kjeldsen A, et al. Diagnostic criteria for hereditary haemorrhagic telangiectasia (Rendu-Osler-Weber syndrome). *American Journal of Medical Genetics* 2000;91:66-7.
230. Sambrook J, Fritsch EF, Maniatis T. *Molecular cloning: a laboratory manual*. 2 ed: Cold Spring Harbor Laboratory Press, 1989.
231. GenBank. National Center for Biotechnology Information, 1999:<http://www.ncbi.nlm.nih.gov/>.
232. Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, et al. The 1993-94 Genethon human genetic linkage map. *Nature Genetics* 1994;7(2 Spec No):246-339.
233. Mukhopadhyay N, Almasy L, Schroeder M, Mulvihill WP, Weeks DE. Manipulation environment for genetic analyses: a data-handling program for facilitating genetic linkage and association analyses: Universities of Pittsburgh, Stanford University, and Carleton College.

234. Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am J Hum Genet* 1996;58(6):1323-37.
235. Alberts MJ, Davis JP, Graffagnino C, McClenny C, Delong D, Granger C, et al. Endoglin gene polymorphism as a risk factor for sporadic intracerebral hemorrhage. *Annals of Neurology* 1997;41(5):683-6.
236. Cymerman U, Vera S, Pece-Barbara N, Bourdeau A, White RI, Jr., Dunn J, et al. Identification of hereditary hemorrhagic telangiectasia type 1 in newborns by protein expression and mutation analysis of endoglin. *Pediatric Research* 2000;47(1):24-35.
237. van 't Hooft FM, Silveira A, Tornvall P, Iliadou A, Ehrenborg E, Eriksson P, et al. Two common functional polymorphisms in the promoter region of the coagulation factor VII gene determining plasma factor VII activity and mass concentration. *Blood* 1999;93(10):3432-41.
238. Girelli D, Russo C, Ferraresi P, Olivieri O, Pinotti M, Friso S, et al. Polymorphisms in the factor VII gene and the risk of myocardial infarction in patients with coronary artery disease. *New England Journal of Medicine* 2000;343(11):774-80.
239. Spek CA, Koster T, Rosendaal FR, Bertina RM, Reitsma PH. Genotypic variation in the promoter region of the protein C gene is associated with plasma protein C levels and thrombotic risk. *Arteriosclerosis, Thrombosis, and Vascular Biology* 1995;15(2):214-8.
240. McGuire W, Hill AV, Allsopp CE, Greenwood BM, Kwiatkowski D. Variation in the TNF-alpha promoter region associated with susceptibility to cerebral malaria. *Nature* 1994;371(6497):508-10.
241. Begbie ME, Chaudhary M, Bernabeu C, Shovlin CL. Functional characterization of a dominant-negative endoglin mutant. *American Journal of Respiratory and Critical Care Medicine*, 2001:A123.
242. Cole SG, Begbie ME, Chaudhary M, Wallace GMF, Shovlin CL. Linkage studies on a PAVM-HHT family unlinked to endoglin or ALK-1. *IVth HHT Scientific Conference: Conference programme and abstracts*, 2001:74.
243. Sarkar G, Sommer SS. Access to a messenger RNA sequence or its protein product is not limited by tissue or species specificity. *Science* 1989;244(4902):331-4.
244. Letamendía A, Lastres P, Botella LM, Raab U, Langa C, Velasco B, et al. Role of endoglin in cellular responses to transforming growth factor-beta. A comparative study with betaglycan. *Journal of Biological Chemistry* 1998;273(49):33011-9.

245. Shovlin CL. Supermodels and disease: insights from the HHT mice. *Journal of Clinical Investigation* 1999;104(10):1335-6.
246. Eassey A, Wallace GMF, Jackson JE, Taylor W, Shovlin CL. Retrospective analysis of the risk of cerebral bleeding in Hereditary Haemorrhagic Telangiectasia. *IVth HHT Scientific: Conference programme and abstracts*, 2001:79.

APPENDIX A – CLINICAL QUESTIONNAIRE

Epistaxis-	when started
	how often
	how serious-?hospital
Telangiectasia-	where
PAVMs-	CXR- when
	?normal
	cyanosis- at rest/on exertion
	haemoptysis
	TIA/CVAs
GI involvement-	upper
	lower
Obs. and gynae.-	effect of periods, pregnancy, menopause
Misc.-	anaemia
	anything else

APPENDIX B - RECIPES

10× Tris Borate EDTA (TBE)

108 grams(g) Tris base

55g Boric acid

40mls 0.5M EDTA pH 8.0

Double distilled water to one litre

Phosphate Buffered Solution (PBS)

1 tablet (made by oxoid) to 100mls double distilled water

Extraction Buffer

70mls double distilled water

5mls 2M Tris pH 8.0

20mls 0.5M EDTA

Genomic Tris EDTA (GTE)

99.3mls double distilled water

0.5mls 2M Tris pH 8.0

0.2mls 0.5M EDTA

3M Sodium Acetate

123g anhydrous sodium acetate

800mls double distilled water- pH to 5.2 with acetic acid then made up to 1 litre
with double distilled water

Agarose Gel Loading Buffer

15mls glycerol

35mls double distilled water

A few grains of bromophenol blue

Denaturing Polyacrylamide Gel Loading Buffer

10mls formamide

10mg xylene cyanol FF

10mg bromophenol blue

200 μ l 0.5M EDTA pH 8.0

APPENDIX C – PRIMERS FOR THE *ENDOGLIN* PROMOTER

Name	Sequence	Melting Temperature (degC)
GW1	TCC.CAA.TCT.CTT.CCT.TCC.TCC.ATG.A	74
GW2	GCC.AGA.GAC.CAG.GGT.TAA.AGT.GAA.A	74
GW3	CTG.GGT.TCA.AAT.CCG.ATT.CTG.CT	68
GW4	GTC.CAA.GGT.CTC.ATG.GCT.AGG.AG	72
GW5	TCT.TCC.TTC.CTC.CAT.GAG.GTG.T	66
GW6	TTA.AAG.TGA.AAG.AGT.CAG.CCC.AGT	68
GW7	GCT.GGA.GTT.GCT.GTC.CGA.AGG.A	70
GW 8	CTC.CAA.TGG.ATG.GCA.GTG.ACA	64
GW 9	CTG.TTA.GGG.TGA.TTG.TTT.TCA.TCT	66
GW 10	CTG.TCC.CCA.TCC.ACG.TTG.TAA.ACA	72
GW 11	GCC.CCA.TGT.CAA.CTG.CAC.TTA.GTA	72
GW 12	CAG.GCT.CTA.GGC.TTG.TCT.CCC.CAA	76
GW 13	TTC.AGT.TTT.CCT.GCC.TGT.TAG.GCA	70
GW 14	GAG.GGT.CAG.GAG.AAG.TGG.ACA.CAG.GGA	86
GW 15	GCG.CAG.ATG.AAG.GCT.GTG.GTA.TGA	74
GW 16	GTT.GTA.AAC.AAG.ACC.CAA.TGG.ATT	66

APPENDIX D – PRIMERS FOR *ENDOGLIN* LINKAGE

Name	Sense primer	Antisense primer
D9S261	AGG.CTG.TGG.TAT.TCC.CAT.T	AGG.TTG.CAA.GAT.GTT.TGA.TT
D9S289	TGC.ATT.AGA.AAA.CCT.TCC.A	CCC.AGC.CTA.TTT.GCA.T
HXB	ATA.GCC.AAA.GAG.AGG.TGC.CC	AGA.GCC.CTT.CTG.TCT.TCT.CC
D9S195	AGC.TCA.GCA.CGG.AGG.G	AGG.GCA.GGT.TCC.TAC.AAA
GSN	CAG.CCA.GCT.TTG.GAG.ACA.AC	TCG.CAA.GCA.TAT.GAC.TGT.AA
D9S60	TTC.TGG.GAA.GCC.TTC.CCT.GA	TGA.ATG.CCA.CTA.TTC.ATG.ATG
D9S315	TTC.TCT.TTG.TCA.TCA.CTT.TTG.G	CTC.AAA.GTA.ACA.CGC.CTG.TG
D9S61	TGA.GAG.ACC.CTC.ACT.CCC.TA	ACA.GAT.GTC.TGT.AGT.CGG.TC
ABL	TTT.ACA.CCT.TCA.CCC.AGA.GA	GGC.TGT.GTT.CAG.TTA.AAC.G
D9S66	CAG.ACC.AGG.AAT.GCA.TGA.AG	CAC.GGG.CAC.ACA.TGT.ATG.C
D9S67	CCA.GGC.TGT.TTA.CTT.TGG.CA	GCC.GGG.AGG.ATG.ACA.TGT.A

Name	Heterozygosity	Size range (base pairs)	Annealing temperature (deg C)
D9S261	0.63	90-104	49
D9S289	0.75	75-87	43
HXB	0.79	107-121	57
D9S195	0.74	204-240	51
GSN	0.70	111-130	51
D9S60	0.81	136-154	53
D9S315	-	~110	55
D9S61	0.80	112-150	55
ABL	0.67	90-117	51
D9S66	0.79	109-135	55
D9S67	0.67	71-113	55

APPENDIX E – PRIMERS FOR *ALK1* LINKAGE

Name	Sense primer	Antisense primer
D12S331	TAA.CAT.TTA.TTC.ATC.TCC.ATA.CTG.A	ACA.TGT.AAG.AGA.GNA.AGT.TTA.CAA.A
D12S85	GCA.CCT.CTC.ACT.CCA.TTA.CA	AAA.TGA.AAG.TCA.AGG.GGA.AC
D12S339	TCA.GGG.TTT.GAG.ACT.NGC.C	GTC.ACG.CCT.GGC.TAG.TGT
D12S347	TTT.TTT.GAG.TCC.AGG.CCA	CCA.GCA.GAG.TTT.TGC.CTT.T
D12S368	GCA.ACA.CCT.TTG.TGA.TGA.AAA.T	AGT.CTG.CAC.AGC.CTG.TCC
D12S325	GGG.CAC.TAT.GGT.GTC.TTC.A	AGG.GCA.TTG.TCT.TAC.TGA.CTG

Name	Heterozygosity	Size range (base pairs)	Annealing temperature (deg C)
D12S331	0.61	175-181	59
D12S85	0.67	105-125	51
D12S339	0.85	257-277	53
D12S347	0.62	96-114	50
D12S368	0.81	200-216	56
D12S325	0.66	209-229	53